DOCUMENT RESUME

ED 044 022 EM 008 346

AUTHOR	Di Vesta, Francis J.; And Others
TITLE	Instructional Strategies: Multivariable Studies of Psychological Processes Related to Instruction. Annual Report, July 1970.
INSTITUTION	Pennsylvania State Univ., University Park. Dept. of Educational Psychology.
SPONS AGENCY	Advanced Research Projects Agency (DOD), Washington, D.C.; Office of Naval Research, Washington, D.C.
PUB DATE	Jul 70
NOTE	262p.
EDRS PRICE	EDRS Price MF-$1.00 HC-$3.20
DESCRIPTORS	*Educational Strategies, *Individual Differences, Learning Characteristics, Learning Motivation, *Learning Theories, Performance Factors, *Psychoeducational Processes, Recall (Psychological), Student Role

ABSTRACT

The studies described in this report were concerned with the effect of active student behaviors, such as note-taking, listening, verbal responding, and test-taking, and the effect of individual differences on the transformation, storage, and retrieval of information. The roles of such individual characteristics as personality, background, and aptitude were explored in two studies. The motivational facets of instruction were the subject of two other studies. Four studies were conducted in an effort to determine the nature of the interaction of instructional treatments with the dynamic properties of individual difference variables. Six studies were completed which relate to the structuring of instructional activities by the teacher and student and to the use of instrumental activities by students. Social-context factors were examined in three studies. (JY)

ED044022

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

INSTRUCTIONAL STRATEGIES:

Multivariable Studies of Psychological Processes Related to Instruction

Francis J. Di Vesta Donald R. Peters Nicholas M. Sanders

Charles B. Schultz Paul D. Weener

Annual Report

July, 1970

Advanced Research Projects Agency

ARPA Order No. 1269

Monitored by

Office of Naval Research

Contract Number N00014-67-A-0385-0006

Department of Educational Psychology

The Pennsylvania State University

EM008 346

+4

ED 044022

ANNUAL REPORT

July 1, 1969 - June 30, 1970

ARPA Order Number:	1269
Program Code Number:	60949
Name of Contractor:	Pennsylvania State University University Park, Pennsylvania 16802
Effective Date of Contract:	1968 September 1
Contract Expiration Date:	1970 August 31
Contract Number:	N00014-67-A-0385-0006
Principal Investigator and Phone Number:	Francis J. Di Vesta 865-6347
Title of Work:	Instructional Strategies

Sponsored by
Advanced Research Projects Agency

ARPA Order No. 1269

TABLE OF CONTENTS

FOREWARD 1

INTRODUCTION AND OVERVIEW 7

An Evolving Theory of Instruction 8

TECHNICAL REPORTS 29

The Effects of Labeling and Articulation on the Attainment of Concrete, Abstract, and Number Concepts 31

Contextual Cues in Cognitive Structures in the Storage and Retrieval of Information 65

The Effects of Presentation Modalities and Modality Preferences on Learning and Recall (Thesis Summary) . . 99

Note-taking and Review in Recognition Learning 107

The Effects of Written Reinforcement and Question Sequence Upon Objective Test Performance 125

The Effects of Studying Together and Grading Procedures on Recall of Subject Matter 157

Reliabilities of Six Personality Measures Used in the Instructional Strategies Research 181

The Effects of Uncertainty, Confidence, and Individual Differences on the Initiation and Direction of Information-seeking Behaviors (Thesis Summary) 197

Recitation Strategies - The Effect of Rates and Schedules of Verbal Responses on Retention 210

Satiation of Divergent and Convergent Thinking and Its Effect on the Need for Novelty (Thesis Summary) . . 229

Small Group Verbal Presentation, Anxiety Level, and Learning 233

PROGRESS REPORTS 249

Project Ikon: Studies of Imagery and Learning 251

Presentation Content, Classroom Notes, and Fact Versus Generalization Learning 257

Note-taking, Rate of Presentation and Immediate Recall . . . 261

Rote and Conceptual Aspects of Classification Learning . . . 265

The Effects of Recall Mode and Recall Interval Expectancies on Recall Performance 269

Foreword

Instructional treatments do not have direct effects on student behavior in a manner that can be adequately conceptualized by a simple input-output model. Students actively confront the material being presented to put it in a form for storage. The variables affecting the student's encounters with the learning task determine, to some extent at least, his motivation for completing the task, which stimuli he will react to, how the material will be transformed, and what parts can be retrieved at a later date. The entire process is further influenced by the overlay of individual differences which interact with teaching methods.

Our current approach has emphasized the active role of the student in acting on the material to be learned; a role which has tended to be ignored in instructional models. In this orientation we have called attention to the importance of student behaviors such as note-taking, listening, verbal responding and test-taking in mediating the transformation, storage and retrieval of information.

The role of the instructor in the instructional process is that of a decision-maker. He sets the stage for learning by structuring the learning situation in terms of some specifiable behavioral objective(s). The decisions he mades are based on principles involving the classes of variables described in this report. On the basis of these considerations, our research to date has emphasized the structure and role of cognitive propensities, cognitive stimulation, the structuring of learning tasks,

the activities in which students engage as learners, and the interactions between or among these variables. An overall view of this orientation is presented in the paper by Di Vesta, entitled, "An Evolving Theory of Instruction," which precedes the research reports.

In these studies we have defined cognitive propensities as filtering agents in cognitive structures. One means of identifying them is by factor analytic studies of self-report measures of personality, achievement, and aptitudes. While the notion of filters may connote, to some, a somewhat static role of individual differences, we would like to emphasize our concern that they be interpreted in the light of a dynamic model of learning as implied by the use of the term "cognitive propensities." Within this area of investigation the following reports have been prepared during the past year:

> The Structure of Selected Personality, Background, and Aptitude Variables Related to Academic Performance. (Sanders, Weener, Di Vesta and Schultz)*
>
> Reliability of Six Personality Measures Used In the Instructional Strategies Research Project. (Sanders and Weener)

The motivational facets of instruction are represented in the present orientation in the form of "cognitive stimulation." This construct emerges as an outgrowth of considerations related to theories based on discrepancy constructs (i.e., doubt, uncertainty, incongruity, or cognitive dissonance). Discrepancy among ideas is assumed to create conditions causing cognitive imbalance thereby goading behavioral or

---

* Asterisked titles were presented as technical reports in the Semi-Annual Report, January 1970 for this contract.

performance changes. As a consequence they lead learners to consider alternatives, to change ideas, or to spend more time in examining new ideas. Two studies related to this idea are:

> The Effects of Uncertainty, Confidence, and Individual Differences on Motivation and Direction-seeking Behaviors. (Schultz)
>
> Satiation of Divergent and Convergent Thinking and Its Effect on the Need for Novelty. (Silvestro)

The dynamic properties of individual difference variables, in interaction with instructional treatments, were assumed to influence the effectiveness of certain stimulus elements, the learning strategies employed by the learner, and the processing of information by the learner. Decisions about instructional strategies, made by the instructor, to parallel this phase of the learning process requires knowledge about modality preferences and how these affect reception learning. There are numerous dispositional variables which need to be considered, most of which undoubtedly remain to be identified along with their behavioral consequences. A modest beginning on the influence of imagery has been made in a study just initiated. It is described in a progress summary, as Project Ikon in the present report. Studies that have been completed on other facets of modality preferences are as follows:

> The Effects of Presentation Modalities and Modality Preferences on Learning and Recall. (Ingersoll)
>
> The Effects of Dogmatism on Learning and Transfer in Concept-Based and Rote-Based Classification Tasks. (Sanders)*
>
> The Effects of Dogmatism in Relation to Expert Endorsement of Beliefs on Problem-Solving. (Schultz and Di Vesta)*
>
> Achievement Anxiety and Performance on the Remote Associates Test. (Weener)*

The storage of information is a critical phase of learning. How and what information is stored are dependent not only on what stimuli become effective stimuli for the learner but also on what form the material takes as a result of the transformation. One can easily imagine, for example, that an experience which is stored only as a picture-image or as an isolated fact will have a different availability for the person than an experience which is stored in the form of a symbolic representation or a generalization, respectively. The transformations employed by the student are affected on the one hand by the way instructors structure the material to be learned and, on the other, by the instrumental activities of the student while studying the material. Completed studies related to the structuring of instructional activities by teacher and student and to the use of instrumental activities by students are as follows:

The Effects of Written Reinforcement and Question Sequence Upon Objective Test Performance. (Peters and Messier)

Contextual Cues in Cognitive Structures in the Storage and Retrieval of Information. (Di Vesta and Ross)

The Effects of Labeling and Articulation on the Attainment of Concrete, Abstract, and Number Concepts. (Di Vesta and Rickards)

Note-taking and Review in Reception Learning. (Peters and Harris)

The Effects of Search Strategies on the Incidental Learning of Concept-Attitudes. (Gray and Di Vesta)*

The Effects of Concept-Instance Contiguity on Concept-Learning. (Sanders, Di Vesta and Gray)*

Learning in any instructional setting involves consideration of the social context. People, or their absence, affect other people's level of anxiety, drive levels, or security. They influence aspirations and

goals. They provide or remove sources of guidance and support. Whatever the direction of its influence the social context certainly affects, for good or for ill, the students expectations and thus, his performance. Often the social context cannot be separated from the student's instrumental activities as, for example, the study related to recitation strategies. However, for convenience the studies which appear to emphasize social-context factors are listed below:

> Small-Group Verbal Presentation, Anxiety Level and Learning. (Weener)
>
> The Effects of Studying Together and Grading Procedures On Recall of Subject Matter. (Sanders)
>
> Recitation Strategies: The Effect of Rates and Schedules of Verbal Responses on Retention. (Schultz)

The current trend in our studies can be seen from the descriptions provided by the progress reports in this publication. It will be apparent to the reader that the view of instruction represented in the completed studies and in the studies in progress is clearly based on a cognitive-perceptual framework. In some instances we made a deliberate attempt to pursue intriguing conceptualizations of learning which have grown out of the current revitalization of cognitive approaches. In most cases we were following our research biases or inclinations by virtue of training or interest. Whatever the reason for following this approach, it appeared to be a fortuitous outgrowth of what initially appeared to be a series of studies related only by an empirical thread and by a common interest in understanding the nature of instruction. In retrospect and prospect it seems to us to be a fruitful one on which to base a viable program of research on instructional strategies.

FDV
NS
CS
PW

INTRODUCTION AND OVERVIEW

An Evolving Theory of Instruction*

Francis J. Di Vesta

An instructional strategy is a metaplan. To paraphrase Miller, Galanter and Pribram (1960, p. 16) it is a hierarchical process employed by the teacher to control the order in which a sequence of operations is to be performed. Thus, the strategy acts as a guide for manipulating stimuli and for transmitting these stimuli in a way that will effectively modify the behavior of another person according to some prestated terminal objective.

Instruction as Communication

The characteristics of the instructional process bear some resemblance to those of the communication process as described by Hovland (1953) and summarized in Figure 1. The plan, which may be compared to a computer program, with its strategies and tactics is stored in the transmitter of the message, whether communicator, instructor, or computer. The flow of the content of the communication, of the arguments or appeals intended to promote attitude change, and of the course content intended to enhance the student's cognitive skills, is channeled, sequenced, structured, and organized according to the plan.

---

* Parts of this article were discussed at weekly seminar meetings attended by Professors Peters, Sanders, and Weener and Dr. Schultz.

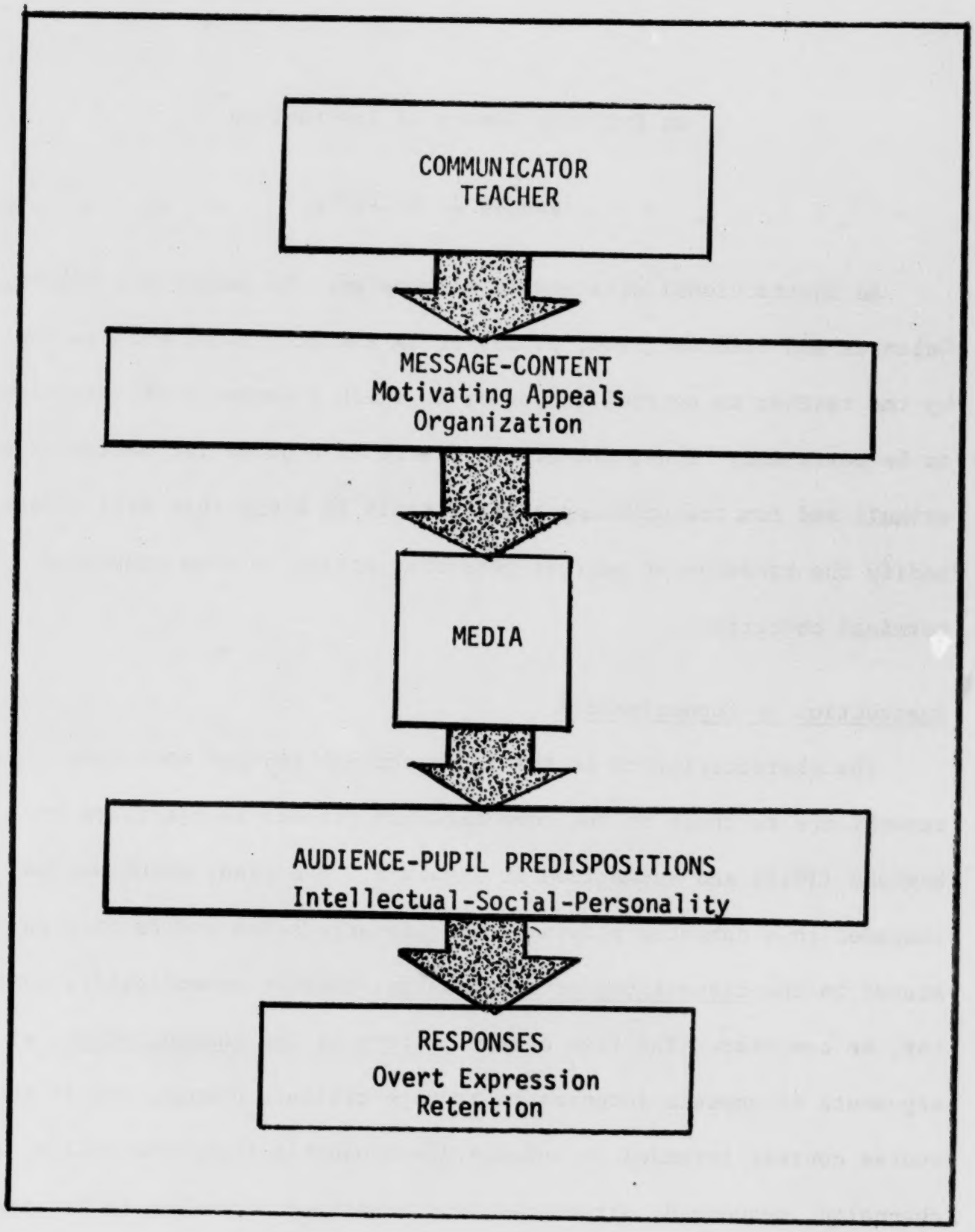


Figure 1. Instruction as communication. This analogue is mainly a convenient basis for classifying variables that influence effective instruction, its main advantage being that all parts are external and observable.

The executive function of the plan governs which of the sub-routines (tactics) will be performed at any one time, thereby providing considerable flexibility in the implementation of the plan from one occasion to the next. The extent to which a message is processed, how it is processed, or even whether processing can or will be attempted depends in large part on the predispositions of the audience or student, that is, on individual differences in social motives, personality factors, and intellectual ability characteristics. The effectiveness of a strategy is determined, and changes within it are made by evaluating the outcomes. In the final analysis, evaluation must always be based, explicitly or implicitly, on the behaviors of the recipient of the communication, that is, the student.

Research within this orientation is typically concerned with the main effects of such conditions as those which belong to the classes of situational, state, and behavioral variables. Accordingly, certain general inferences or hypotheses about the instructional process become apparent and immediately available as topics for educational research. Thus, for examples: The personality of the instructor ... his trustworthiness, and his expertise ... and the cues he provides or the lack of them (as for example, in computer-assisted instruction) can influence the acceptance of a communication. Implicit in the communication content is it's ability to arouse motivation or uncertainty in the recipient. Material logically or psychologically sequenced; arranged in hierarchical fashion on the basis of end-products of learning or on the basis of intellectual skills (Ausubel, 1968; Gagne, 1970); or presented in a motor, ikonic, or symbolic mode will make decidely different contributions to the end-products of learning. Information about these topics

should ultimately feed back into the instructional process to affect decisions that must be made as a part of the instructional strategy.

A Model for Research on Learning and Instruction

This general orientation can be extended as indicated in the original proposal, and later with some elaboration in the semi-annual report (January, 1970), by incorporating the interactions between and among these variables into the research program. Perhaps the single most widely publicized of these interactions, at the present time, is the so-called aptitude by treatment interaction (ATI) implying that instructional methods are most efficient when matched with individual differences whether in the form of personality or intellectual variables. Walberg (1970) suggests a model very similar to that described here, with, perhaps, somewhat more emphasis on environment, though instructional variables must be included by definition. His formulation, as does the present one, asks such questions as (Walberg, 1970, p. 187):

1. Which instruction best promotes learning?
   (f_1 = summative evaluation.)
2. Which students learn best?
   (f_2 = studies of prediction and selection.)
3. Which environments best promote learning?
   (f_3 = stimulation and enrichment.)

A model representing the relationships among these variables and of their interactions are summarized in the following equations (walberg, 1970):

$$L_h = f(I_i, A_j, E_k)$$

$$L_h = f_1(I_i)+f_2(A_j)+f_3(E_k)+f_4(I_iA_j)+f_5(I_iE_k)+f_6(A_jE_k)+f_7(I_iA_jE_k).$$

Aptitude by Treatment Interactions

In our earlier statement of the aptitude by treatment interaction, which specified a relatively straightforward functional relationship, only the behaviors of the student in response to the task were considered in a description of the dependent variables (i.e., criterion performance). Further consideration of this point suggested that certain instructional and study activities must also be brought into the model and thereby raised another series of questions related to decisions an instructor must make, as follows (p. 6 - semi-annual report, this project):

1. What is it that students do while the instructor is "instructing?"
2. What activites do students engage in between the time of onset of instruction and the elicitation of the criterial or terminal performance? How do these activities affect performance?
3. If such student behaviors are important to learning, what can the instructor do to manipulate such behaviors to maximize performance?

These questions tended to place the research emphasis on student activities which affect processing for storage and retrieval of information. They brought to the fore note-taking, verbal responding (e.g., directed student response, self-verbalization, and verbalization to peers) and test-taking as major instrumental activities. These instrumentations were viewed as having two roles in the student's behavior: They could be seen as possible terminal activities (for example, instructional variables can and do affect the kind of notes students take or the kinds of study activities they engage in before taking tests); as mediating activities which transform performance characteristics ordinarily elicited by given instructional variables (for example, the student who prepares for a multiple-choice examination probably achieves quite different

objectives than one who studies for an essay examination). In either role, these activities could be modified by aptitudes and or could modify further the influence of aptitudes on performance characteristics. Thus, it can be seen, that the student's instrumental activities may be considered as independent variables, as mediating variables, or as dependent variables influenced by and being influenced by aptitudes or individual differences.

While this approach was a fruitful one, in the sense of generating a number of studies on variables related to instructional strategies (semi-annual report, 1970), it was a relatively static model. A critical examination of it called attention to the dynamic properties of learning which were noticeable by their absence. As a consequence of this orientation, instructional variables were now viewed as processes used by the instructor to set the stage for learning; aptitudes were seen as readiness patterns which act as filters permitting the learner to benefit by certain environmental-instructional conditions but also to be hindered by others; instrumental activities were translated into transformational mechanisms aimed at processing information for storage and retrieval; and learning criteria now encompassed not only achievements and end-products but also abilities represented in the application, use, and retrieval of information

A Dynamic Model of Learning and Instruction

Since the writing of the semi-annual report, the latter notions about the characteristics of the learning-teaching situation have been extended into an even more detialed description of the learning process as it appears to function in an instructional setting. An attempt at a

dynamic approach appears at this juncture to be more useful for guiding research than does the model previously described.

The present model is an evolving one. Accordingly, the presentation here must be considered as tentative. Whether the order of the stages and other details are accurate must be determined by further investigation. Nevertheless, the model, for the present, can serve as a means of summarizing the research reported here, can point to variables which enter into decisions that eventually become a part of instructional strategy, and can point to areas which require further investigation. While, for the most part, the description here is of the dynamics of the learning process with occasional reference to instruction, the ultimate description should indicate parallel activites by the instructor.

An overview. The major stages that must be considered by the instructor are outlined in Figure 2. Briefly, this sketch acknowledges an input by the teacher and output in the form of student behavior. Furthermore, it considers the social context within which the instructional process occurs. While these three classes of variables are ostensibly open to direct observation, the appearance is deceptive since the meanings of these variables, in the last analysis, must be implied.

Between the input and output are two major stages which can only be inferred. Nevertheless, they suggest a highly active, adapting, and dynamic organism since they suggest ways in which instructional materials are processed by the student. In the first stage, attending and perceiving are required for an analysis of the input. Individual differences (filters) determine whether the stimuli are or can be potentially meaningful ones. If not, there is further analysis provided

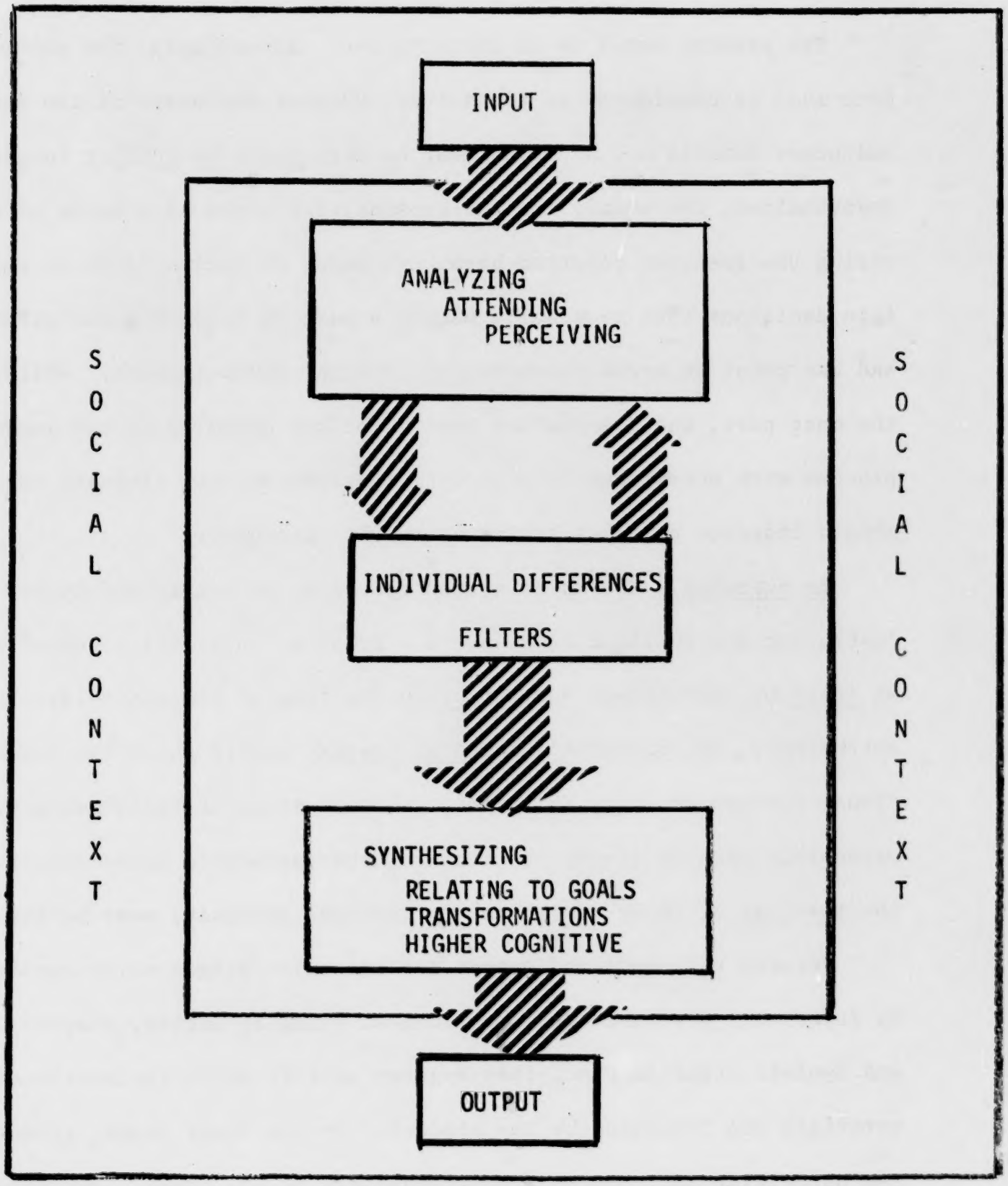


Figure 2. An overview of a model based on a dynamic view of instruction.

the student is motivated to continue. If he is no longer motivated he would exit (literally or figuratively) from the learning situation.

Once particular stimuli are selected they are subjected to further processing for storage and retrieval in the synthesis stage. At this point, instructional materials take on interpretations which are idiosyncratic to the learner. Motivations, too, change character for they now seem to be peculiarly cognitive or epistemic in quality. Such notions as incongruity, dissonance, curiosity, uncertainty, and imbalance are employed to indicate that motivation is derived by a perceived discrepancy between the learner's present state and his anticipated state of achievement.

Transformation of the instructional material, however, is the principal processing that goes on during the synthesis stage. It can be as simple as mere association of the new material with a mnemonic device (as in the "30 days hath September" ... rhyme) or it can be as complex as integrating vast bodies of knowledge into a formula comprised of less then a half dozen symbols (e.g., $E = mc^2$). Whatever the transformation, the key word appears to be coding, the understanding of which may also be the key to the understanding of the higher mental processes.

The analysis stage. The details of the first stage of processing by the learner are depicted in Figure 3. The input phase is entirely under the control of the instructor. What he does, and the decisions he makes at this point depends on his theory of instruction. The elements of this phase are essentially the same as those presented in the communication model. Research programs dealing only with this phase would be directed solely toward investigations of the effects of treatments. Accordingly, the main concern would be with the direct

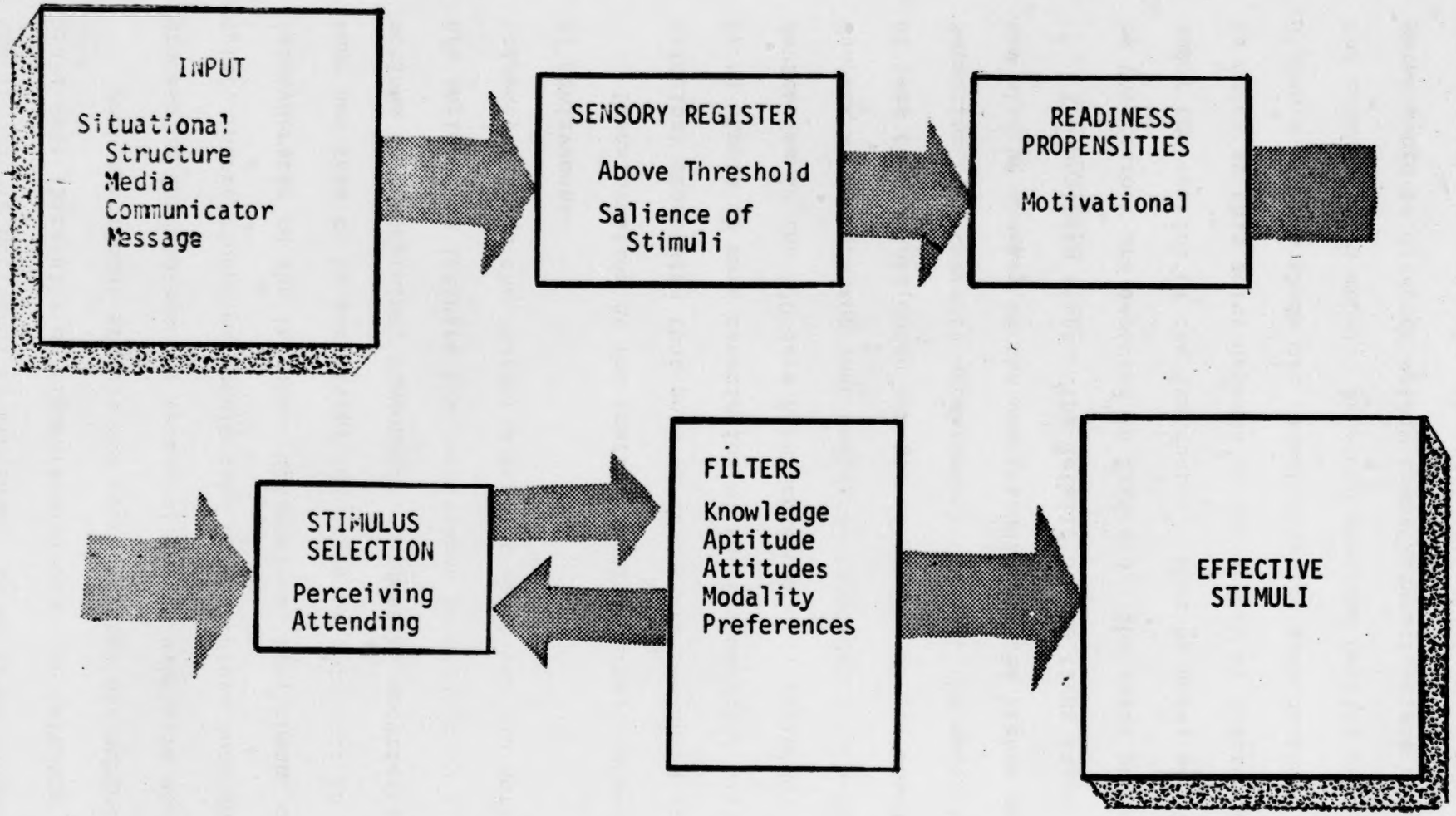


Figure 3. The analysis stage in processing materials in instruction and learning.

effects on student performance of such variables as sequencing of subject matter, types of advance organizers, modes of presentation, contextual cues, task difficulty, and characteristics of the instructor all of which are external to the student. An important feature of the present analysis is the recognition that whatever occurs at this point in instruction can only provide potential stimuli for the student. Oftentimes these are classified as nominal stimuli.

Before the stimuli from the input become effective there must be a considerable amount of preliminary processing. Initially, the message and accompanying stimuli must be registered. Accordingly, they must, at the least, be above threshold and salient to the learner. With this condition met, a degree of readiness in the form of a learning-set (e.g., curiosity or the need for achievement) provides the motivation for perceiving and attending; a process which culminates in focal attention. This means that all the features of a given situation are not automatic elicitors of behavior. More likely they are optional. Which structural features are attended to, and the method of analysis employed, differ from person to person.

The features that are selected by different observers or by the same observer at different times are assumed to be, in large part, a function of the filter-system, which is comprised of all so-called individual differences variables. As an illustration, differences in acquired knowledges or aptitudes differentially determine the effective stimuli. If the stimuli cannot be analyzed, they do not become effective. Recycling may be necessary between the filter and the perceptual-attending system until a pattern is constructed. The exact characteristics of the pattern are left unspecified but they may emerge

as figure-ground or as meaningful dimensions. Because different features are selected for attention, analysis is a constructive act. Thus, there will be considerable variability, among students in a class, in what they observe even though they experience the same input.

Effective stimuli. The effective stimuli, or constructed pattern, result from the attentive-perceptive mechanisms. They comprise the common link between the analysis and synthesis stage. Under carefully prescribed environmental conditions, such as those that are obtained in classical-conditioning laboratories, the behavior predicted from the input would closely approximate that predicted from the effective stimuli; maximum differences would be obtained when the input is highly ambiguous. In general, the less-prescribed the external controls the more opportunity there will be for idiosyncratic selections of stimuli from which configural patterns will be formed. The notion of effective stimuli includes the idea of "interpretation of the situation" thereby taking into account the phenomenological experiences of the student in the learning situation. (The relationship between the effective stimuli and interpretation should, probably, be represented by a link or, perhaps, by a feedback loop in the diagram.) The interpretation is that part of the effective stimulus pattern which is comprised of task demands as implied from the task itself or from instructions; goal expectations which result from prior experiences and are therefore influenced by the filter system; and processing strategy preferences. Thus, the effective configural pattern to which the student reacts is comprised of selected stimuli from course material or course content and of expectations regarding desired outcomes. The incorporation of expectations into this

part of the model appears especially important to explain differences that occur among students in the kinds of transformations they use.

The synthesis stage. A student in a learning situation has at least two behavioral alternatives during the analysis stage: either exits from the situation or he processes the information. In the latter alternative certain features of the input are selected, as already described. Then, in the synthesis stage, these stimuli are put into a perspective consonant with his interpretations of the learning situation (i.e., What is expected of him by the instructor? How long is the material to be retained? What kinds of goals are to be achieved? and so on). Once this point has been reached the input is encoded; it is categorized, (which may require nothing more than recognition of the item), elaborated, or otherwise synthesized. What is synthesized need not be clear or distinct as already noted. It is the synthesis that contributes to clarity. (See Figure 4.)

How the input is synthesized, or the extent to which it is synthesized, depends in large part on the student's expectations (interpretations). These appear to direct further processing of the input as part of their executive function. Expectationa may be several forms: Task demands can be implied from instructions, from assignments, from the demand characteristics of an experiment, and from characteristics of the task (e.g., problem-solving vs. memorizing a poem). Goal expectancies relate one's performance to the criteria or standard characterizing the terminal performance. They may range from the desire to reach a high standard of excellence by the student with high need for achievement or satisfaction with a mediocre performance by students with low need

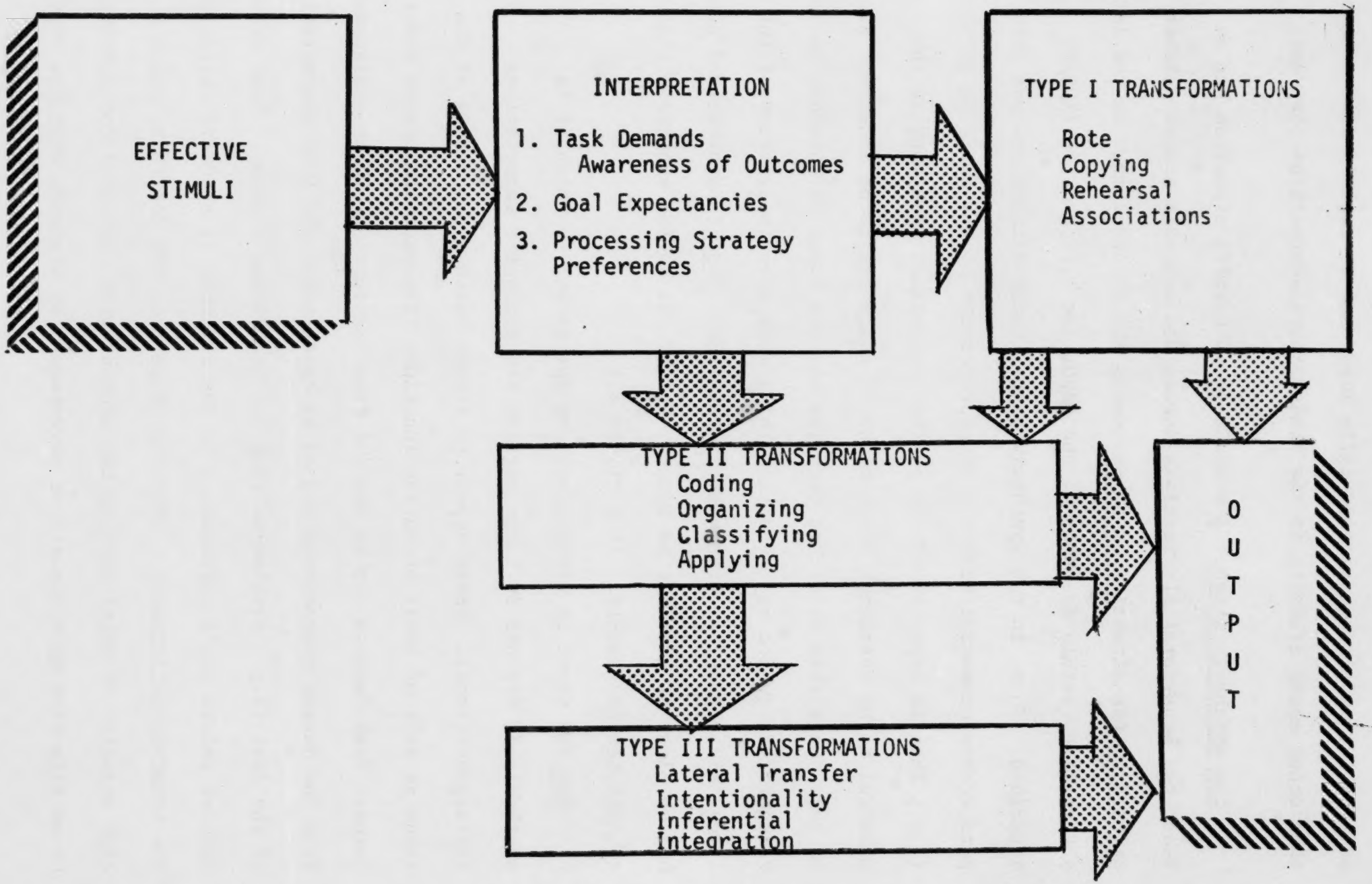


Figure 4. The synthesizing and storage phases of instruction and learners.

for achievement. Students with previous experiences of success may try to reach realistically higher goals than previously; those with previous experiences of failure may set unrealistically high or low goals. Goal expectancies may be imparted directly to the student when he is instructed on such matters as the kinds of tests he will be given, or when he is given certain kinds of advance organizers, or when certain grading provisions are specified. They are also influenced by the social context in which learning occurs, and by the normative standards of one's peers or peer group. Finally, expectancies can be affected by learned preferences for one learning strategy over another. Thus, a student who succeeds at rote memorization may view all tasks as being most successfully approached through rote memory while another student may try to encode all materials in meaningful ways.

Interpretation, as it is being employed here, always involves the weighing of what must be done with the material against the criterion to be reached. By this definition, interpretation determines what will be done with the materials. A wide range of instrumental activities may be employed for reconstructing the effective stimuli into patterns that will implement the goal activities suggested by the interpretations. All essential processing activities in this phase are related to transformation of the incoming stimuli. For convenience in the present account, the kinds of transformations have been classified at three levels, and are presumed to be arranged hierarchically according to complexity. This arrangement implies the desirability of sequencing instruction in ways that parallel these kinds of transformations. The aim served by the transformation is to store the material in a form that will lessen memory load and that will make it available for later retrieval.

The transformation at Level I are relatively primitive. For convenience, the transformations at this level are called associative because they appear to consist mainly of arbitrary associations within the material itself (for example, linking one sentence to another). In general, the modification bears some resemblance to the new learning or at least is only a step away from the new learning as, for example, they might be in a free association task. The student predisposed to process material at this level may attempt to memorize materials on rote, verbatim, or arbitrary bases; he may attempt to form some elementary images of the material; or he may make some relatively low level associations. These processes are similar to those used in "cramming" for example, where the student may expect to take a test requiring only recall, to retain the material for only a brief period of time or where he will be satisfied with minimal achievements. It should be noted that students whose interpretations require more advanced levels of transformations probably must master Level I transformations first. Overlearning, repetition, practice, rehearsal, and copying are important instrumental activities at Level I if the student is to master information, to retain it, and to protect it against interference. Retrieval of information here is typically of the recall or recognition variety. Interference (i.e., retroactive and proactive inhibitions) is its greatest enemy.

Level II transformations involve attempts to make the material meaningful. These are <u>constructive</u> transformations. Modifications at Level II are similar to the content of experience only on an abstract dimension. The most typical example of Level II transformations is concept-formation. In principle, these transformations code the material

in a form that approximates existing cognitive structure. They are constructive in the sense that new organizations (for the student) of ideas are often achieved. Thus, for example, the learner may organize the new learning in terms of existing concepts, he may acquire a new classification (concept), or he may find an application for the learning. The instrumental activities for constructive transformations are encoding according to arbitrary mnemonic systems (the very lowest level), encoding according to thematic schemes, encoding in terms of existing cognitive structures, classifying what is learned, and organizing material in logically sequenced ways. Retrieval of information at this level is dependent on cues that aid in identifying the correct plan or "storage area."

Level III transformations are <u>inventive</u>. As a class they comprise the epitome of the higher mental processes. These transformation represent a major leap from the form of the original learning experience and often bear no resemblance to it. In lateral transfer, for example, the person generalizes over a broad set of situations at the same level of complexity as he would when learning the relation between two sides of a right triangle and transferring it when seeing, for the first time, a problem in physics relating to acceleration of a body rolling down an inclined plane (Gagne, 1970, p. 335). Characteristic of Level III transformation is the testing of alternatives to arrive at unique implications or unique organizations of material already acquired by the learner. Included at this level are such behaviors as the identification of new relationships among concepts (i.e., principle-formation) and the identification of a unique solution to a problem. Hence, we speak here of intentionality, inferential processing, integration, and

restructuring. Level III transformations, at the highest level of development, must be considered integrative, inventive, productive and constructive. The learner at this level engages in behaviors which emerge as novel sequences and which are reproduced in easily communicable plans comprised of clearly defined hierarchical arrangements of behavioral units.

Output

Ideally, the behavioral output will reflect the expectations of the learner and the transformations he employs. There are numerous possibilities that might be enumerated here but will not be because they have not been developed sufficiently. Others are omitted because they require further exploration. However, it can be noted briefly that output may be defined in terms of type of test (e.g., recall or recognition); kind of end-product (e.g., motor-skill, attitude, or concept); kind of intellectual skill (e.g., learning-to-learn, learning-to-perceive, or learning to test the alternatives); or in terms of the characteristics of the terminal performance (e.g., fast or slow, or higher or lower, than previous performance). Which of these is used by the instructor or investigator will be determined by the decision about what is to be tapped ... the effects of selective perception? of expectations? or of transformations?

Epilogue

The model presented here and the considerations it highlights points to a sort of hierarchy of learning processes including attending, perceiving, discriminating, selecting, and transforming. All of these are processes assumed to be essential facets of the learner's activities. Further elaboration of this model will require: specification of stages

that can be influenced by instruction and the kinds of instructional activities that are required to facilitate learning at each of these stages; a more complete specification, than is currently available, of the kinds and characteristics of instrumental activities in which the learner can engage at each stage of learning to reach specified terminal objectives; and a more detailed specification of the kind of outcomes than can be expected at each of the phases described above. Some progress has been made in each of these areas but further elaboration must depend upon additional empirical evidence.

References

Ausubel, D. P. Educational psychology: A cognitive view. New York: Holt, Rinehart & Winston, 1968.

Gagne, R. M. The conditions of learning. Second Edition. New York; Holt, Rinehart & Winston, 1970.

Hovland, C. I., Janis, I. L., & Kelley, H. H. Communication and persuasion. New Haven: Yale University Press, 1953.

Miller, G. A., Galanter, E. E., & Pribram, K. H. Plans and the structure of behavior. New York: Henry Holt, 1960.

Walberg, H. J. A model for research on instruction. School Review, 1970, 78, 185-199.

TECHNICAL REPORTS

Summary

The Effects of Labeling and Articulation on the Attainment of Concrete, Abstract and Number Concepts

Francis J. Di Vesta and John P. Rickards

Technical Problem

This study makes the assumption that learning is mostly verbal and conceptual. Accordingly, it was hypothesized that if labels were imposed on learning materials at a different conceptual level than was required by the terminal criterion of performance these labels would interfere with performance even though they were entirely accurate. Conversely, labels congruent with the conceptual level required for terminal performance were expected to facilitate performance. Since previous findings regarding the role of articulation (overt verbalization) have been inconclusive, this variable was also manipulated. The expectation being that if learning did, indeed, require verbal transformation, overt verbalization of labels would enhance the beneficial or detrimental effects of labeling. In addition, it appeared that the Remote Associates Test might be a measure of verbal encoding ability and therefore should interact with the treatments described above.

General Methodology

The treatments were administered experimentally in a laboratory setting. The stimuli were presented via a projector. The task was to associate several, all different objects (which were labeled by S as described above) with a commonly shared name in the form of a novel

monosyllable. The only way the task could be learned was by the process of conceptualization.

Technical Results

The results were as follows: Concrete concepts were learned more rapidly than abstract concepts which in turn were learned more rapidly than number concepts. Labels that were too highly specific or too highly generalized hindered performance while labels that represented a particular conceptual level facilitated performance. Furthermore, the main effects of labeling (i.e., hindrance or facilitation) were significantly increased under the full articulation requirement compared to the partial articulation requirement. None of the main effects interacted with the Remote Associate Test Scores.

Educational Implications

The results of this experiment imply an order of "readiness" for learning concepts which should be considered when presenting new material. Thus, concrete illustrations probably should precede more abstract formulations. Symbolic materials in mathematical form appear to be more difficult to grasp, i.e., they take longer to learn. Suggesting to the student, in advance of the learning task, some overview of the material appears to be a desirable practice. However, it is possible to cast this overview at a conceptual level that will mislead the student. Accordingly, instructors should be sensitive to the terminal performance they will require of their students when employing "advance organizers." Since verbalization appears to fix an idea more firmly it may be advisable to require a verbalized answer to a question only when there is some certainty that the student will be able to culminate his reply with the desired response.

The Effects of Labeling and Articulation on the Attainment of Concrete, Abstract and Number Concepts

Francis J. Di Vesta and John P. Rickards

In a classic paper, Heidbreder (1946) described a cleverly conceived investigation on concept-formation. In brief, the task required that S respond with a nonsense syllable, via the anticipation method, to each of the pictorial stimuli in a list. The unique characteristic of the experiment was such that the stimuli from one block of trials to the next were always dissimilar. However, they were conceptually related according to the qualities of object, shape, and number. As a result of these relationships among lists, it was possible for the S to learn common responses to items in all lists. Heidbreder found that the concept of object was learned more easily than the concept of shape which in turn was more easily learned than the concept of number. She concluded, "...the perception of concrete objects is the dominant mode of cognitive reaction" (p. 214). The results implied that the process of learning the concept may be explained via the use of such constructs as mediating responses (e.g., Kendler and Kendler, 1962) or hypothesis-testing (see for example, Bourne, 1968).

There is some difference of opinion regarding the manner in which the mediating mechanism is supposed to function. Some (e.g., Bousfield, 1961) think of it as a chain of competing responses. Accordingly, in a covert naming process, the object initiates a range of associates or

selection of names that could be applied to it. Osgood (1961), on the other hand, suggested that mediation occurs via a mechanism of placing the object within a continuous semantic space of meaningfulness. The Kendlers (Kendler and Kendler, 1962), taking still another position, considered the mediator to be a response that directs the attention of the perceiver or learner to a dimension of the stimulus. The analysis of these positions has failed to generate experiments which clarify the theoretical interpretations of the nature of the mediating response (Hunt, 1962). Nevertheless, it is probably correct to assume that visual stimuli are encoded verbally (Neisser, 1967) and that inhibiting the occurrence of the correct mediator should hinder concept learning. Under optimal conditions the subject can locate the set of attributes or dimensions elicited by the exemplars, and can make the appropriate discriminations, provided the dimensions are employed as cues for the naming response. During learning, then, the task is one of making the relevant cues and required response contiguous. "The stimuli produced by the mediating response become decision criteria of the concept of the name. If they can be associated with an object, that object may be assigned the name" (Hunt, 1962, p. 80).

In addition to the processes described above, which for purposes of brevity may be classified as the labeling function of naming, there is the question of the means by which the labels or names can be produced. Thus, the person may not be aware of the production of the mediator, he may "think of it" in a very vague sort of way, he may "say it to himself" in very specific terms, or he may articulate it overtly. A name that is articulated overtly commits the learner to a selection of the mediator. If the selection is "correct," learning should be

facilitated. If incorrect, learning should be hindered. When names are articulated covertly, or where the mediator is present but in a vague or ambiguous form, correspondingly greater degrees of flexibility for correcting a wrong response, after feedback, are provided.

On the basis of the above rationale it was hypothesized that the degree of overt or covert labeling of incoming stimuli affects the dimensions that are perceived and selected and, consequently, affects the rapidity with which the concept can be acquired. Certain assumptions underlie this hypothesis. First, it was assumed that the S's "thinking" can be channeled through control of the coding processes by instructions (Gagne, 1970). Second, when the subject is instructed to code in a given way, that code (label or name) is as likely to compete, as it is to be congruent with the learner's subjective code. Such competition might take many alternate forms. For example, the learner might prefer to code objects first; however, if the experimental manipulations forced him to code numbers first, learning would be impaired. Similarly, coding a picture of a face as an object rather than as belonging to the class of people might conflict with the learner's subjective code thereby interfering with acquisition of the code. Comparable activities are probably frequent occurrences in the classroom and other everyday situations. In these settings it would be expected that learning would be most rapid where the subjective and normative codes coincide. Learning would be least rapid where subjective and normative codes are antagonistic and thereby compete to create interference. Third, for some tasks at least, labeling can be varied along a continuum of specifity with the concept typically lying somewhere between the label for the specific object at one extreme and the label for the highly generalized

category at the other extreme. The latter category ordinarily represents a degree of generalization beyond that required in the concept-formation task.

Instructions to label and even instructions to provide specific labels are frequently employed in concept-learning tasks without recognition of their effects on the demand characteristics of an experiment. These manipulations may imply that some objectives (e.g., serial-order learning) are to be achieved to the exclusion of other objectives (e.g., classification of items). Accordingly, since a concept-learning task requires conceptualization, the learner who is instructed to label the specific items at the concrete level (e.g., a man's face or a pine tree) will be at a disadvantage; he enters the task at a nonconceptual level and will be working with too many items. The learner who labels the incoming stimuli according to some scheme (code) that reflects the experimenter's code will perform in maximal fashion; he enters the task at the precise level of abstraction required in the experiment. Finally, the learner who is instructed to label at a level of abstraction beyond that required by the task will be at a disadvantage; he will be working with too few categories. Nevertheless, the latter instruction does have the advantage of allowing the learner to infer that he is to perform at a conceptual level. It interferes with optimal learning to the extent that the learner must proceed to "breakdown" the superordinate concept into other classifications before he can reach criterion.

While casual observation suggests an influence of overt verbalization (i.e., articulation) on learning, the evidence for either a positive or negative influence on paired-associate learning is inconclusive (Underwood, 1964; Di Vesta & Ingersoll, 1969). Gagne and Smith

(1962), on the other hand, found that verbalization of moves in a problem-solving task facilitated the learner's ability to arrive at a solution. There is some evidence, too, that overtly verbalized labels are retained better than nonverbalized labels (Carmean & Weir, 1967). This effect may be due to the increased amount of time the learner attends to an articulated label and to the possibility that auditory stimuli are stored more easily than visual stimuli.

If articulation has an effect on retention, as suggested, the effect may be either facilitative or disruptive, depending upon the materials to be learned (Weir & Helgoe, 1968). Accordingly, it was hypothesized that articulation interacts with the level of generalization represented by the labels or names. More explicitly, articulation would result in poorer performance than nonarticulation when the learner labels items with either a concrete (i.e., specific) name or a superordinate category name. In either case, the saliency of the items is increased through articulation; it thereby tends to impede the acquisition of the concept. Conversely, when the label is appropriate, articulation has a beneficial effect; the saliency of the code, already at the optimal level of generalization, is enhanced and thereby would facilitate concept acquisition.

The present experiment was designed to investigate the hypothesized effects of labeling and articulation on concept acquisition. The task and overall paradigm employed was similar in all essential characteristics to the one described by Heidbreder (1946). Thus, it was also possible to reexamine the order in which the three types of concepts were learned.

Method

Design

The Ss in this experiment learned to label exemplars of three classes of objects under different labeling and verbalization conditions. The anticipation method of presenting paired-associates was used in the presentation of the task. The stimuli were drawings of objects and the responses were nonsense words. In most respects the stimuli, objects, concepts, and responses resembled those described by Heidbreder (1946). A total of 11 unique lists were used for each S. The labeling variable was comprised of three levels. At one level the instructions implied that each drawing depicted a specific, independent object; at a second level the instructions implied that each drawing represented a class of objects; and at a third level the instructions implied that each drawing represented one of three general concepts: Object, shape or number. These conditions were orthogonally crossed with two levels of instructions to verbalize or to articulate the name of the object: in one variation the S verbalized the name of the object at the appropriate level of generalization (i.e., physical object, particular concept, general concept) for the condition to which he was assigned; in another set of conditions the S overtly verbalized the names of objects at the appropriate level of generalization for some of the lists (i.e., a set of nine drawing-nonsense word pairs) but did not verbalize the names for one-third of the lists. The Ss were administered lists until they reached a criterion of one completely correct trial. The basic overall design implied a 2 x 3 factorial analysis of variance.

Subjects

The Ss were 60 college sophomores enrolled in an introductory educational psychology course at The Pennsylvania State University. Although participation in the experiment was voluntary the Ss received credit toward their course grade for such participation. Each S was assigned to one of the conditions within a block of six treatment-combinations (n=10) by reference to a table of random numbers. None of the Ss had participated previously in a concept learning task.

Materials

The stimulus materials were a modification of Heidbreder's (1946) materials. The stimulus lists consisted of drawings of objects paired with one syllable, four-letter nonsense words as responses. Each drawing was an exemplar of a concept. In each series there were nine drawings, and each drawing had a different nonsense word paired with it. In successive series the nonsense words representing a particular concept class remained the same, but the drawings representing exemplars of the concept class were changed. The nine concepts represented in any one series could be classified into three superordinate concepts. These stimuli and associated response terms are classified in the list presented in Table 1. The main differences between the lists for the present experiment and those used by Heidbreder were as follows: Actual physical objects were always represented in the drawings; the "face" concept was replaced with an "animal" concept and the "tree" concept was replaced with a "vegetable" concept. Examples of stimuli are displayed in Figure 1.

Table 1

A Listing of the Labels at Each of the Three Levels of the Labeling Condition and of Their Associated Responses

General Concept	Particular Concept	Physical Objects	Response
Object	Animal	bear, camel, elephant, cat, giraffe, squirrel, dog, house, lion, and rabbit.	RELK
Object	Vegetable	asparagus, radishes, mushroom, cucumber, carrot, pepper, pumpkin, corn, peas, and squash.	MULP
Object	Building	teepee, log cabin, church, igloo, farmhouse, windmill, castle, house, skyscraper, and lighthouse.	LETH
Shape	Circle	flower, drum, clock, coin, balloon, wheel, ring, tennis ball, wreath, and globe.	FARD
Shape	Loop	snake, fishing rod, train tracks, belt, arrow, tie, chain, rope, necklace, and wire.	STOD
Shape	Crossed Pattern	shovels, twigs, swords, rolling pins, ski poles, pencils, brooms, cattails, flags, and canes.	PRAN
Number	Two	chairs, sleighs, telephones, guitars, cactuses, stockings, shoes, hats, books, and angels.	LING
Number	Five	snowmen, lamps, cups, dollar signs, spoons, anchors, bells, candles, cards, and ice cream cones.	DILT
Number	Six	sailboats, baskets, trees, pipes, umbrellas, ants, fish, leaves, bottles, and musical notes.	MANK

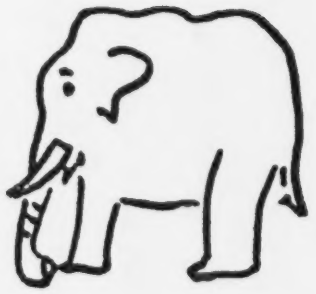
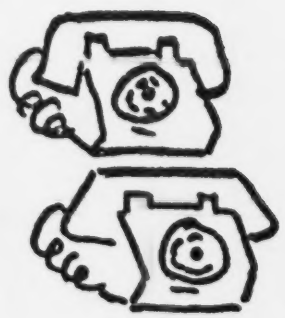
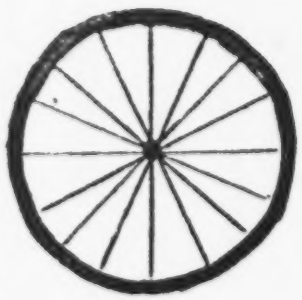
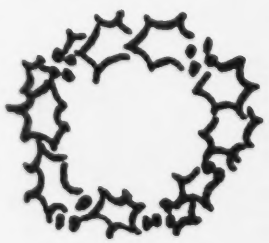
List 1	List 2	List 3	List 4	Response Terms
				RELK
				LING
				FARD

Figure 1. Examples of stimuli and associated response terms for the animal, twoness, and circle concepts employed in four experimental lists.

Within a single series, the nine drawings were arranged according to Heidbreder's (1946) rules which were as follows (p. 180-181): "(a) Each third of the series contained an instance of one concept of each of the three categories - one instance of a concept of a concrete object, one of a concept of a spatial form, one of a concept of a number; (b) no instance was followed by an instance of a concept belonging to its own category - e.g., no drawing representing a concept of a number was followed by one representing another concept of a number, but it might be followed by one representing a concept either of a concrete object or of a spatial form; (c) from series to series, the order within a series was varied so that no position was occupied with more than chance frequency by instances of one concept, so that no regular sequences occurred, and so that possibly advantageous positions, such as first and last in the series, were distributed equally with respect to the nine concepts."

There were ninety-nine different drawings in all (i.e., eleven series). One series was used for pretraining purposes, and the rest constituted the training series. Another ten series were generated from the ninety drawings of the first ten training series. The pictures and orders within each of the second ten lists were as dissimilar as possible from the pictures and orders within any of the first set of ten lists. The same within-series rules were employed in the development of the second set of ten lists as were used in the original set of lists. All stimuli and stimulus-response pairs were photographed for presentation via a Dunning Animatic Projector.

Procedure

The S and E were seated at opposite ends of a table, 9-ft. in length. The stimulus (drawing) and stimulus-response (i.e., drawing-nonsense word) pairs were rear-projected onto a transluscent screen directly in front of S. The anticipation method was utilized at a 3:3 seconds presentation interval with a 6 second rest interval after each block of nine trials.

Each S was first given standard paired-associate instructions, in which he was informed of the nature of the learning task. These and all subsequent instructions were read to S by the E. After this introductory phase, E then read the instructions to S appropriate for the particular condition to which he had been randomly assigned. Following this, the training series was given. The experiment was terminated when S reached a criterion of nine correct anticipations of the concepts in any one series. Any S who did not have more than a total of five correct anticipations out of the first ten series was dropped from the experiment. Instructions to induce the conditions of the experiment were administered immediately prior to the pretraining series.

Labeling conditions. The essentials of all labeling conditions are outlined in Table 1.

In the physical object conditions, Ss were instructed that each drawing depicted some concrete object and that they were to name each drawing with its particular concrete object name as soon as the picture appeared on the screen. For example, when a drawing of a bear appeared on the screen, the Ss were to respond with the label "bear" and then to respond with its new (nonsense word) label.

In the particular concept condition, each S was instructed that while each drawing depicted an object, it "could also be classified in a more general way," that is, each drawing represented some concept and so, Ss in this condition were instructed to name the particular concept that each drawing represented as soon as it was presented to them. For example, when a drawing of a bear appeared on the screen, the Ss were to respond with the concept "animal," after which they responded with its new (nonsense word) label.

The Ss in the general concept condition were told that an object can be classified on many different levels of generality, and that each of the drawings shown represented one of three concepts - shape, number, or object. They were further instructed that a drawing belonged to the number category if it consisted of more than a single object, and if each of the objects was separate from the other(s). Secondly, the object in a drawing represented a shape concept if form stood out or seemed to predominate in the drawing. And thirdly, the Ss were told that a drawing belonged to the object category if there was only one object depicted, or if there was more than one, they were collectively one. Finally, Ss in this condition were instructed to name the superordinate concept to which each drawing belonged as each was presented to them and before they responded with the new label. For example, when a drawing of a bear appeared on the screen, the Ss were to respond with the superordinate concept "object" and then to say the new label.

In all three conditions, the S was provided with examples appropriate to the instructions. The first list was a list to aid the S in implementing the instructions.

Articulation conditions. Half of the Ss in each of the previously mentioned conditions were instructed to articulate the labels appropriate for their particular labeling condition as each drawing appeared on the screen before them. In the partial articulation condition, each S was informed that for some of the lists, he would not be required to articulate the labels appropriate for his particular condition. That is, during this time S was free to use any labeling system he chose or none at all and he need only say aloud the nonsense word appropriate for a given drawing. The partial articulation conditions were so-called because S did not articulate the labels for the drawings in every third list beginning with the second list. He articulated the labels for the objects depicted on each of the remaining lists.

The scoring was done by E during the experiment. It required a check mark for every correct response and a "zero" for every incorrect response. Subsequent to the experiment E asked S some questions about the experimental experience.

Results

Number of Correct Responses

The number of correct responses were analyzed by a mixed analysis of variance, having two between and two within factors. The between factors were the two levels of the Articulation variable and the three variations of the Labeling condition; the within factors were the three Kinds of Concepts and six Blocks of Trials with two trials in each block.

A summary of this analysis is presented in Table 2. The cell means for all experimental conditions are summarized in Table 3. In

Table 2

Summary of Analysis of Variance for Number of

Correct Responses Over Six Blocks of Two Trials

Source	df	MS	F
Between subjects			
Articulation (A)	1	12.25	1.31
Labeling (B)	2	396.54	42.47 ***
A X B	2	79.51	8.52 ***
Error between	54	9.34	
Within subjects			
Trials (C)	5	375.83	286.71 ***
A X C	5	4.64	3.54 **
B X C	10	12.43	9.48 ***
A X B X C	10	1.91	1.45
Error within	270	1.31	
Kind of Concept (D)	2	60.61	16.06 ***
A X D	2	1.60	.42
B X D	4	29.46	7.81 ***
A X B X D	4	5.49	1.45
Error within	108	3.77	
C X D	10	3.39	3.69 ***
A X C X D	10	.63	.68
B X C X D	20	2.12	2.30 *
A X B X C X D	20	.61	.67
Error within	540	.92	

*** $p < .001$

** $p < .01$

* $p < .025$

Table 3

Mean Numbers of Correct Responses and Mean Numbers of Trials to Criterion for all Experimental Conditions

Kind of Label and Concept	Articulation Treatment			
	Complete		Partial	
	Number of Correct Responses	Trials to Criterion	Number of Correct Responses	Trials to Criterion
Physical Object Label				
Object Concept	2.62	9.20	2.67	9.80
Shape Concept	1.02	17.10	1.77	14.10
Number Concept	.90	18.00	.60	17.70
Particular Concept Label				
Object Concept	4.32	6.20	3.20	8.60
Shape Concept	4.17	6.50	2.68	9.90
Number Concept	4.27	6.60	3.02	9.90
General Concept Label				
Object Concept	2.18	10.90	2.32	10.30
Shape Concept	1.88	13.50	2.43	10.70
Number Concept	1.48	13.70	2.23	12.20

brief, the results were as follows: The main effect due to Articulation was not significant ($F < 1.00$). The main effect due to Labeling yielded F (2,54) = 42.47, $p < .001$. The order of difficulty of learning under the various labeling conditions (from easiest to hardest) was as follows: Particular Concept label ($\bar{X} = 3.61$), General Concept label ($\bar{X} = 2.09$), and Physical Object label ($\bar{X} = 1.59$). The Articulation X Labeling interaction yielded F (2,54) = 8.52, $p < .001$. The means representing the interaction are graphically displayed in Figure 2. As shown in this graph, the Ss in the Articulation condition performed better than Ss in the Partial Articulation condition for only the particular concept level of the three labeling conditions.

As would be expected the main effect due to Blocks of Trials was significant, yielding F (5,270) = 286.71, $p < .001$. As shown in the graph in Figure 3, the results of the analysis also yielded F (5,270) = 3.54, $p < .01$, for the Blocks of Trials X Articulation interaction. Here it can be seen that by the eleventh trial the Ss in the Partial Articulation condition are performing better than those in the Articulation condition. The effect due to Labeling X Blocks of Trials yielded F (5,270) = 9.48, $p < .001$. This interaction is depicted in Figure 4.

The main effect due to Kind of Concept yielded F (2,108) = 16.01, $p < .001$. This finding implies a clear replication of Heidbreder's results. That is, collapsing across conditions, the order of difficulty of the various concepts (from easiest to hardest) was the following: Object Concepts ($\bar{X} = 2.88$), Shape Concepts ($\bar{X} = 2.33$) and then Number Concepts ($\bar{X} = 2.01$). The interaction between Kind of Concept X Labeling yielded F (2,108) = 7.81, $p < .001$. The graphic presentation

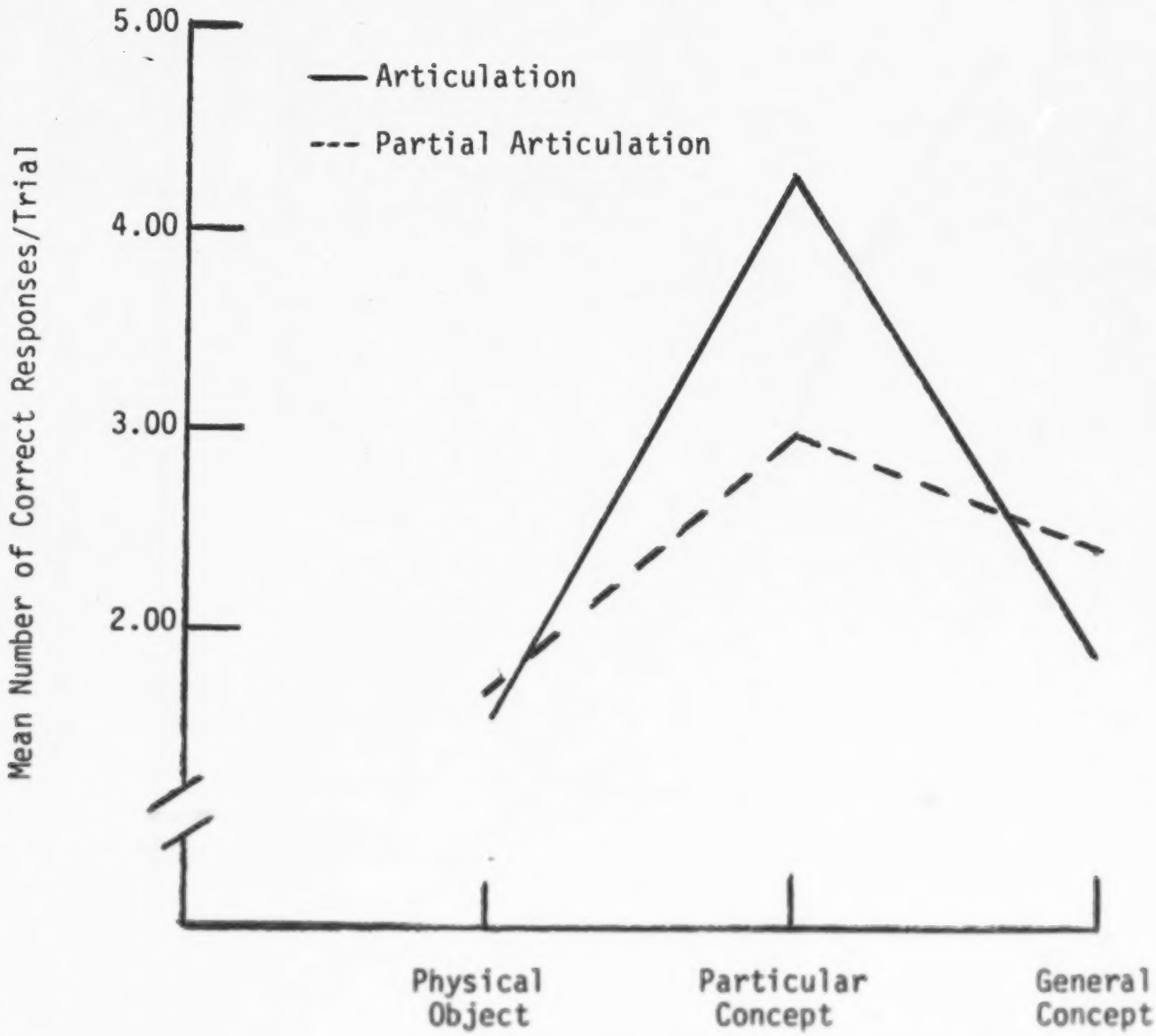


Figure 2. Mean number of correct responses (totaled over 12 trials) as functions of Labeling and Articulation.

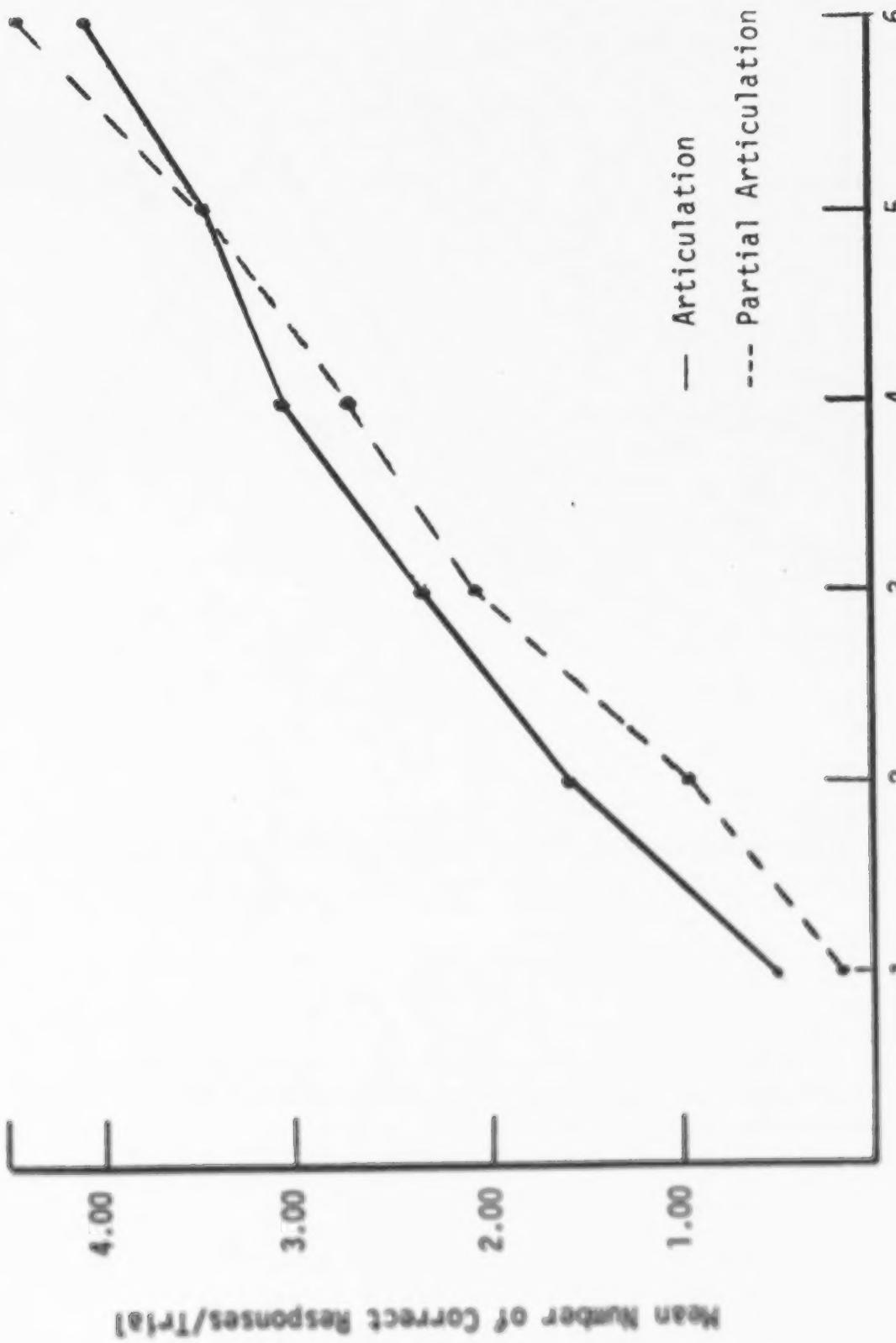


Figure 3. Mean number of correct responses over six blocks of trials as a function of articulation instructions.

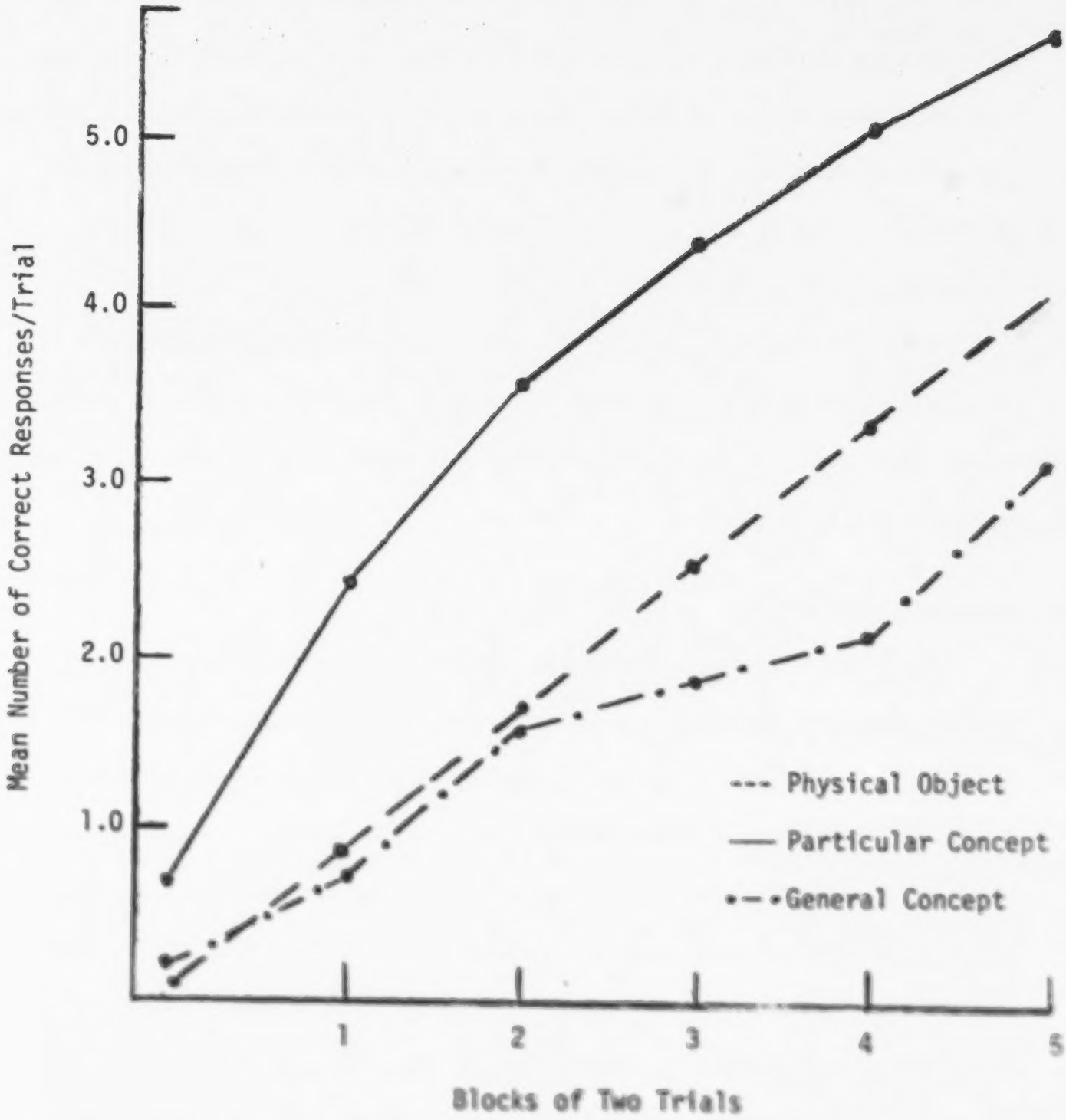


Figure 4. Mean number of correct responses over six blocks of trials as a function of labeling conditions.

of this interaction in Figure 5 illustrates that while Ss in the Particular Concept label condition are relatively unaffected by the kind of concept to be identified, the Ss in the Physical Object label condition perform quite differently with respect to the various kinds of concepts they were to identify. Further, the trend of the General Concept label condition is in the same direction as the Physical Object labeling condition. The Blocks of Trials X Kind of Concept yielded F (10,540) = 3.67, $p < .001$. The means for this interaction are summarized in Figure 6.

Finally, the second order interaction of Labeling X Blocks of Trials X Kind of Concept, which yielded F (20,540) = 2.30 was significant ($p < .025$). None of the other main effects or interactions was significant ($p > .05$).

Trials to Criterion

The trials to criterion (i.e., the first trial in which every instance of any general concept was correctly identified) were analyzed by a mixed analysis of variance, having two between and one within factors. The between factors were the two levels of the Articulation variable, and the three levels of the Labeling condition; the within factor was Kind of Concepts. A summary of this analysis is presented in Table 4. A summary of the cell means for the various levels of the conditions represented in this analysis is presented in Table 3. In all major respects the results of this analysis were comparable to those in the analyses of numbers of correct responses described above.

The effect due to Articulation was not significant ($F < 1.00$). The main effect due to Labeling yielded an F (2,54) = 23.74, $p < .001$.

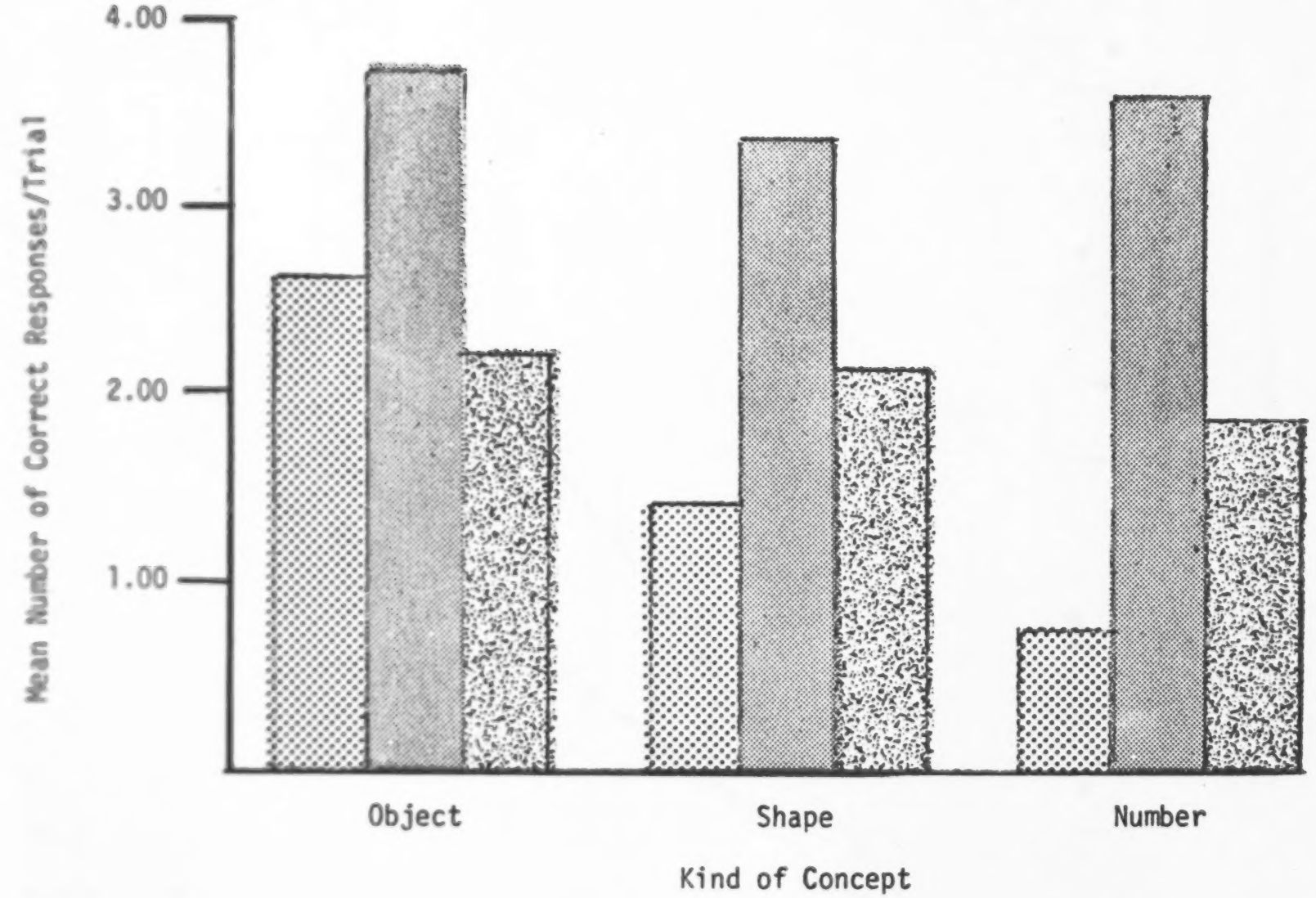


Figure 5. The mean number of correct responses in attaining three kinds of concepts under different labeling instructions.

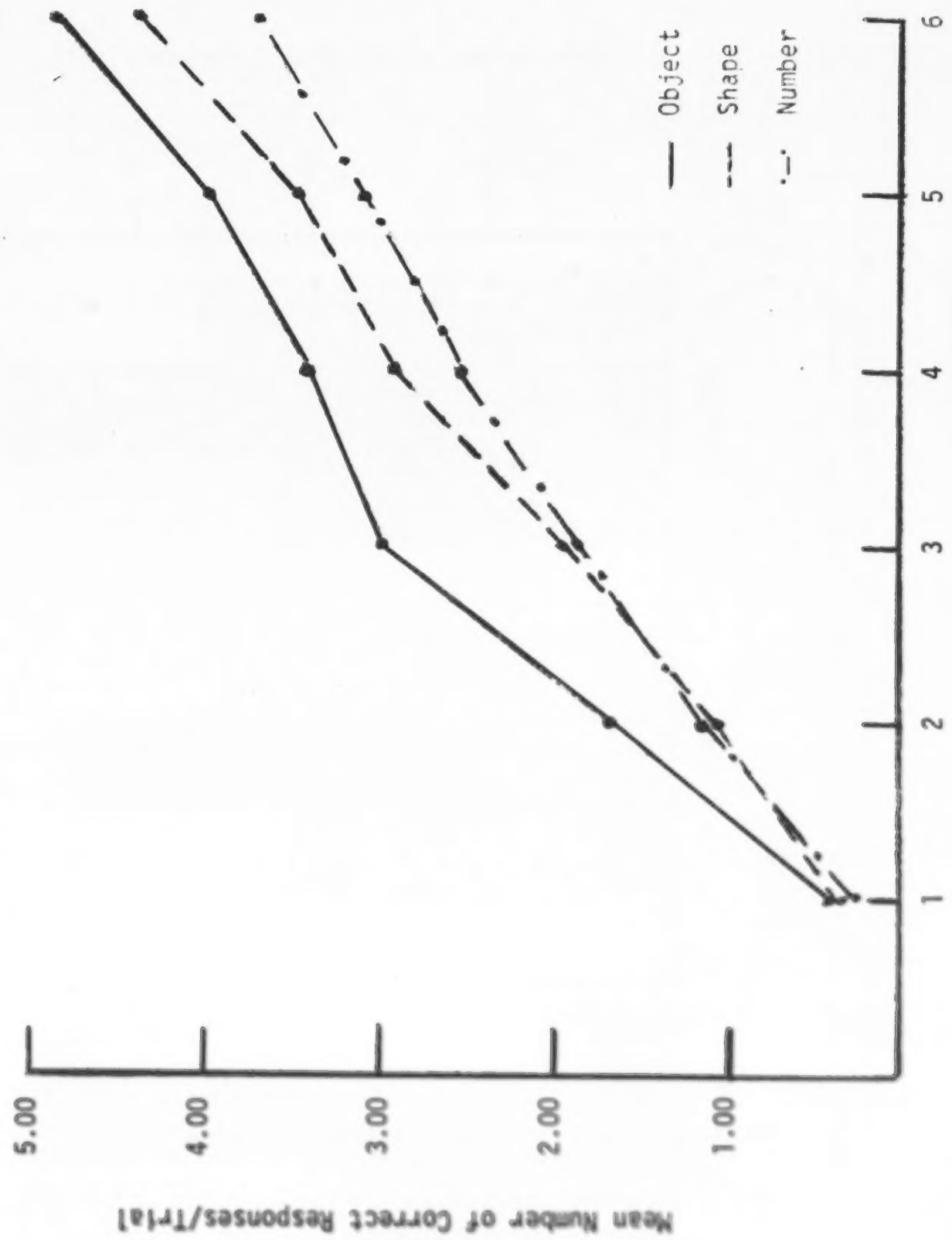


Figure 6. Mean number of correct responses over blocks of trials for each of the three kinds of concepts.

Table 4

Summary of Analysis of Variance

for Trials to Criterion

Source	df	MS	F
Between subjects			
Articulation (A)	1	1.25	.05
Labeling (B)	2	619.27	23.74 **
A X B	2	94.47	3.62 *
Error between	54	26.08	
Within subjects			
Kind of Concept (C)	2	237.65	25.15 **
A X C	2	10.85	1.15
B X C	4	84.27	8.92 **
A X B X C	4	7.17	.76
Error	108	9.45	

** $p < .001$

* $p < .05$

Collapsing across conditions the means for the three levels of the Labeling variable were as follows: Physical Object label ($\overline{X}$ = 14.32), Particular Concept label ($\overline{X}$ = 7.95), and General Concept label ($\overline{X}$ = 11.88).

The interaction of Articulation X Labeling yielded F (2,54) = 3.62, $p < .05$. The nature of the interaction was identical to that found in analyses of numbers of correct responses. (See Figure 2).

The effect due to Kind of Concept, F (2,108) = 25.15 was also significant ($p < .001$). Collapsing across conditions the mean number of trials to learn the three kinds of concepts were: Object Concept ($\overline{X}$ = 9.17), Shape Concept ($\overline{X}$ = 11.97) and Number Concept ($\overline{X}$ = 13.02). Further, there was a significant Labeling X Kind of Concept interaction which yielded an F (4,108) = 8.92, $p < .001$. Figure 5 of the previous analysis is illustrative of this Labeling X Kind of Concept interaction in this analysis. The second order interaction was not significant ($p < .05$ in this analysis.

Discussion

Order of Concept Attainment

The results of this study provided a clear replication of Heidbreder's (1946) study which in all practical respects involved only minor changes in stimuli and procedures. Thus, the Ss first learned concrete concepts, then abstract concepts, and finally, number concepts. Heidbreder's explanation was within the Gestalt frame of reference. As indicated by Hunt (1962), "the dominance of mediating responses (e.g., 'that's a pair of things') which was associated with a name was determined by the natural tendency to perceive concrete,

familiar objects (good Gestalt) without abstracting smaller stimulus features. Therefore, the concepts based on objects should be easiest to learn to use, then concepts based on physically present 'part qualities,' the patterns, and finally concepts based on the abstract number aspect" (p. 127).

An alternative explanation of these results was provided by Baum (1954) who argued that it was not the object-like quality that determined the order in which the concepts were learned but, rather, it was the processes involved in Gibson's (1940) generalization-discrimination hypothesis. Later, Grant and Curran (1953) using analytically defined stimuli on the Wisconsin Card-Sorting Task obtained the same order of attainment as in the present study when the stimuli were presented at random on the cards. However, when the geometric stimuli were arranged in orderly fashion number concepts were easiest to learn. In explaining these results, Hunt (1962) indicated "... this is because the regular arrangement of forms creates a stimulus pattern that is perfectly correlated to number. Subjects respond to the overall pattern and not to the more abstract concept of number. If the forms are regularly arranged on the card only the leftmost (rightmost) boundary of the pattern need be established to determine the number of forms on the card. If the forms are irregularly arranged, the location of each figure must be established, as well as its separate identity. This means that the boundary, not of the pattern but of the separate figures, is important. Grant and Curran's results are consistent with [an] analysis of dimensions and values based on stimulus scanning" (p. 129).

The Effect of Labeling

The labeling condition to which an S is assigned functions to establish an "expectancy" related to the task demands. This expectancy, in turn, influences the stimuli to which the S attends. Thus, the Ss in this experiment were clearly affected by the level of generalization of the label they were required to use. Labels that were too specific or too highly generalized hindered performance when compared with the use of the particular-concept label, thereby supporting the original hypothesis.

The present findings are in accordance with Bruner's (1956, 1957) analysis which suggests, as summarized by Hunt (1962), "Perception is viewed as an act of inferring wholes from usually valid cues obtained from parts of the stimulus. ... The first step in perception is a primitive categorization of the stimulus by identifying a set of possible percepts. This set can be used as the basis of future 'guesses' about object identity. After each guess, specific tests can be carried out to validate it" (p. 129). More particularly the use of labeling is very much like putting the stimuli, from which the concept was to be learned, into a highly overlearned template. The ones which introduce the most "noise" interfere the most with concept-attainment. Conversely, the template which is related to the structure of the concept, the template matching scheme which accentuates the information to be abstracted from the stimulus is also the one which facilitates the attainment of the concept.

The explanation presented immediately above, is not unlike that provided by Neisser (1967), who suggests a cognitive analogue of the perceptual processes of "focal attention" and "figural synthesis"

(Neisser, p. 300). Accordingly, attention is the allocation of cognitive resources to a certain part of the field (of attention); the attentive process is determined at least in part by the expectancies of the task to be performed. The aspects of a stimulus to which the person will attend is partly determined by his expectations. Not only is focal attention determined by expectancy, but figural synthesis is also determined in part by this nonstimulus variable (Neisser, p. 103 and 301). Thus, both analysis (the stimulus features to which an S will attend) and synthesis (the configural pattern that he will "construct" or synthesize) is influenced by S's expectation. For example, Neisser (p. 59), citing research by Bruner and Minturn (1955), notes that a stimulus is identified as "13" when the S is expecting numbers, but becomes "B" when the S is expecting letters.

The Effects of Articulation

The results related to the effects of the articulation conditions in the present experiment supported the second hypothesis. They implied that the main effects due to labeling were enhanced under the articulation requirement. Thus, compared to partial articulation, overt verbalization of all responses resulted in greater interference associated with the use of the "physical object" label and of the "general concept" label thereby tending to result in depressed performance in both cases; whereas the use of the "particular concept" label clearly enhanced performance.

There are several alternative explanations of this effect. Carmean and Weir (1967), for example, speculate that verbalization may have an influence on the relative time spent in attending to the stimuli, may

result in putting the learning material into a form (e.g., symbolic) that can be stored more readily than another form (e.g., pictorial), or may have "special" consequences that enhance memorial processes. In actuality, the second alternative above suggests a possible reason why verbalization had an effect in this experiment but did not in other experiments. Accordingly, in the present experiment while the stimuli were presented pictorially, articulation had the effect of emphasizing the transformation of these pictures into a symbolic (verbal) form that could be readily stored. On the other hand, in Underwood's (1964) and Di Vesta and Ingersoll's (1969) experiment articulation had no effect becuase only verbal stimuli were used. In other words, the stimuli were already in a form that could be stored, and there was no special advantage to be gained from articulation. This explanation also seems to coincide with Underwood's (1964) suggestion that articulation may influence the recall process by somehow changing the structure of the unit when it was put into memory storage. Gagne and Smith (1962), who found that verbalization facilitated problem-solving provided an explanation that was somewhat more ambiguous but probably means something similar to that provided by Underwood. They say that "... the content of the verbalizing during practice was fairly pedestrian and to some extent routine, so that it could be readily categorized. What then accounts for its effect on problem-solving? In answering this question we have no theory to call upon. It would appear that requiring verbalization somehow 'forced the Ss to think.' In other words, this treatment may have had the effect of constantly prodding the Ss to think of new reasons for their moves ..." (Gagne & Smith, 1962, p. 17).

Summary

Sensory perception of concreteness appears to be more rapid than that of abstractness or number. This order of attainment may be due to dominance, preference, or complexity associated with the stimulus materials; or, it may be due to the utilization of some, as yet unidentified, order of searching for a given cue. Despite the alternative explanations provided in the discussion, the evidence concerning how subjects structure tasks is sparse. Whatever clues are to be derived from further experimentation will undoubtedly come from more precise statements of perceptual theories than are presently available.

The expectancies that were established via the labeling conditions clearly imply an influence on the attention and scanning processes. When the label designating the "physical object" was used, the S was misdirected in the sense that he was led to expect something other than a conceptual task. Accordingly, they attended to and synthesized the learning materials in a different way (and in a debilitating manner) than did Ss who were assigned to either of the conceptual levels of the labeling condition.

Finally, our explanations of the articulation condition correspond to those provided by Neisser (1967) and Gibson (1969) both of whom theorized that visual information is often recoded in verbal form and then stored verbally in memory. Support for this view comes from Conrad (1964) who found that substitution errors in immediate recall tend to involve units that sound alike, even when the original stimuli are visual. Thus, it would be expected, as was found in the present experiment, that instructions to code pictorial stimuli in a verbal form appropriate to the task requirements would not only avoid inter-

ference with the processing of stimuli but, in fact, may facilitate learning. By the same reasoning, an articulated verbal code that is discrepant from the appropriate code would interfere with verbal coding processes that might ordinarily be employed by the learner.

References

Baum, M. Simple concept learning as a function of intra-list generalization. Journal of Experimental Psychology, 1954, 47, 89-94.

Bourne, L. Jr. Concept attainment. In T. R. Dixon and D. L. Horton (Eds.), Verbal behavior and general behavior theory. Englewood Cliffs: Prentice-Hall, 1968.

Bousfield, W. A. The problem of meaning in verbal learning. In C. N. Cofer (Ed.), Verbal learning and verbal behavior. New York: McGraw-Hill, 1961.

Bruner, J. S. Going beyond the information given. In Contemporary studies of cognition. Cambridge: Harvard University Press, 1956.

Bruner, J. S. On perceptual readiness. Psychological Review, 1957, 64, 123-152.

Bruner, J. S. & Minturn, A. L. Perceptual identification and perceptual organization. Journal of General Psychology, 1955, 53, 21-28.

Carmean, S. L. & Weir, M. W. Effects of verbalizations on discrimination learning and retention. Journal of Verbal Learning and Verbal Behavior, 1967, 6, 545-550.

Conrad, R. Acoustic confusions in memory. British Journal of Psychology, 1964, 55, 75-83.

Di Vesta, F. J. & Ingersoll, G. M. Influence of pronounceability, articulation, and test mode on paired-associate learning by the study-recall procedure. Journal of Experimental Psychology, 1969, 79, 104-108.

Gagné, R. M. Problem solving. In A. Melton (Ed.), Categories of human learning. New York: Academic Press, 1964.

Gagné, R. M. The conditions of learning. Second Edition. New York: Holt, Rinehart & Winston, Inc., 1970.

Gagné, R. M. & Smith, E. C. Jr. A study of the effects of verbalization on problem solving. Journal of Experimental Psychology, 1962, 63, 12-18.

Gibson, E. J. A systematic application of the concepts of generalization and differentiation to verbal learning. Psychological Review, 1940, 47, 196-229.

Gibson, E. J. Principles of perceptual learning and development. New York: Appleton-Century-Crofts, 1969.

Grant, D. A. & Curran, J. F. Relative difficulty of number, form, and color concepts of a weigl-type problem using unsystematic number cards. Journal of Experimental Psychology, 1953, 43, 408-413.

Heidbreder, E. The attainment of concepts. I. Terminology and methods. Journal of General Psychology, 1946, 35, 191-223.

Hunt, E. B. Concept learning: An information-processing problem. New York: John Wiley & Sons, 1962.

Kendler, H. H. & Kendler, T. S. Vertical and horizontal processes in problem-solving. Psychological Review, 1962, 69, 1-16.

Neisser, U. Cognitive psychology. New York: Appleton-Century-Crofts, 1967.

Osgood, C. E. Comments on Professor Bousfield's paper. In C. N. Cofer (Ed.), Verbal learning and verbal behavior. New York: McGraw-Hill, 1961.

Underwood, B. J. Articulation in verbal learning. Journal of Verbal Learning and Verbal Behavior, 1964, 3, 146-149.

Weir, M. & Helgoe, R. S. Vocalization during discrimination: Effects of a mixture of two types of verbalization patterns. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 842-844.

Contextual Cues and Cognitive Structures in the Storage and Retrieval of Information

Francis J. Di Vesta and Steven Ross

Technical Problem

Learning materials are presented and studied within some contextual arrangement, if meaningfully learned; or in the absence of a contextual arrangement, if arbitrarily learned. In particular, this investigation was based on the assumption that context leads to a change in structure or patterns that enter into transformations of learning material by the learner. The purpose of the present study was to understand the ways in which verbal contexts can be manipulated and the ways in which these variations affect learning, storage of information, and retrieval.

General Methodology

The task was presented in a laboratory setting and was administered individually. It consisted of two phases: learning and transfer. Within each phase there were 20 word-pairs to be learned. One word of each word-pair was supplemented by two accessory words for the purpose of manipulating the contextual variable. In the transfer task the focal word (i.e., the one associated with accessory words in the learning phase) was replaced by either one of the original accessory words, by the concept represented in the contextual configuration, by another concept for which the word might be an exemplar, or by a word

unrelated to the context or to the focal word. The primary measure was the number of correct responses on the first trial of the learning phase and the first trial of the transfer phase.

Technical Results

One finding of this study was that meaningful contexts actually slow down the initial learning of adult Ss. It was assumed that this finding implied a dynamic process in which adult Ss attempted to find the meaning of the arrangement even though they were not instructed to do so. On the other hand, a meaningless context was rejected and time was spent in learning the word-pairs by arbitrary (relatively so) means. The slower rate of initial learning within the meaningful context did not prove to be a handicap in the transfer phase on related materials. However, it was clearly the case that the context, incidentally learned, did affect transfer to other contexts.

Educational Implications

Context during learning, whether provided by the instructor or imposed by the learner, is related to transfer and retention. A meaningful context may require more study on the part of the learner but it also has greater payoff in transferability to related materials than does an arbitrary context. In addition, there is more material learned when a meaningful context is provided, in the sense that the context is learned incidentally. The disadvantage is that the meaningful context may restrict or delimit the possibilities for transfer by inducing a set of expectancy that the material can only be applied in a limited way. This would mean that an instructor should provide a number of contexts if greatest transfer is to be achieved.

Implications for Further Research

An important dimension to be investigated is the relationship between single- and multiple-contexts during learning on later transfer to a range of applications. In addition, this study is linked to Project Icon, described elsewhere in this annual report, in the sense that materials may be presented in a visual or verbal context. Presumably, learners oriented to learning via imagery will learn more about and from the visual context than would those oriented toward learning via symbolic material. Finally, the effect of context, in the form of advance organizers, behavioral objectives, and the like, on the learning of text-like prose appears to be a logical extension of the present study.

Contextual Cues and Cognitive Structures in the Storage and Retrieval of Information

Francis J. Di Vesta and Steven Ross

This study views man as a cognizing organism who brings order out of the otherwise chaotic bombardment of stimuli to which he is subjected. He forms rules, he categorizes, he organizes, he patterns, he codes, and he classifies these external events . . . whether or not he is directed to do so by an outside agent. There is now sufficient evidence to indicate that these are generalized tendencies of intellectually mature persons . . . even of immature ones. An understanding of what is acquired, what is stored, and what is retrieved requires an understanding of the process of pattern recognition, the ways in which events are encoded, or stored, and the ways in which they are decoded, or retrieved.

The present study is based on several assumptions underlying this process, as follows:

1. Experiences are stored in memory as idiosyncratic (subjective) patterns. The more ambiguous an experience, the more unique (i.e., private) will be the code (pattern) by which it is stored. On the other hand, to the extent that codes are shared by members of the language (as a basis for symbolizing) community, the more apparent (predictable) will be the coding of an experience by members of the community.

2. Both contextual cues, associated with an experience at the time of storage, and the person's experiential history (cognitive structure) determine the final form the pattern takes . . . i.e. the way in which it is coded.

3. Efficient retrieval of an experience requires that the person have access to, or recognize, the same pattern (i.e. code) via contextual cues by which the experience was initially stored.

4. Two or more experiences embedded in similar contexts will be more difficult to retrieve individually than when they are embedded in different contexts. Similarly, when the experience is embedded in a context where discriminable characteristics are much like those of the experience itself then the experience becomes part of the abstracted pattern. It is sacrificed to the more general pattern even though its topographical features differ from those of the contextual cues. Lacking discriminability, on the bases of relevant identifying features, it will be unavailable on later occasions for retrieval.

Although a number of hypotheses are suggested by the above rationale, the present investigation examines the notion that fortuitous (or adventitious) backgrounds have significant effects on focal items and may themselves be learned or, in some way affect learning. Take, as an illustration, the word "club." When it is placed in the context of gun, sword, and knife, its potentiality for transfer and the situational requirements for its retrieval seem to be entirely different than when it is placed in the context of group, band, boy scouts, and people. It is the purpose of this experiment to understand these affects with verbal stimuli which comprise a large part of the instructional stimuli. Furthermore, instruction can be presented

arbitrarily or it can provide for a context. If the latter, it can provide for one of a number of contexts. Presumably, which context is provided can affect learning but probably the greatest effect of context is on retrieval and transfer. In particular, then, this investigation was directed by the assumption that context leads to a change in structure or patterns that enter into transformations by the learner and it is the pattern that is stored, retrieved, or transferred.

Experiment I

Design

The experimental design consisted of three orthogonally crossed variables in a 2 x 2 x 2 factorial design with repeated measures. Each S within a specific condition was required to learn all pairs in an initial list of twenty paired-associates. Half of the Ss received lists in which one accessory word was positioned above and another below the stimulus term; and the other half studied a list of paired-associates identical to the first list in all respects except that the stimulus and response terms were reversed. Thus, in the first list the stimulus term was bounded above and below by the accessory words and in the second list the response term was bounded by the accessory words. The second manipulated variable dealt with the meaningfulness of the imposed context as defined by the relationship of the accessory words to the stimulus or response term to which they were proximately located. Thus, in one set of conditions the context words were meaningfully related to each other and to the specific term suggesting an inclusive concept category; in the other conditions the context

words were unrelated either to each other or to the paired-associate term. The groups were further subdivided during the transfer task which followed the initial learning trials. The treatments consisted of replacing the stimulus or response term which had been within a contextual framework by either a) one of the two previously given accessory words or b) by the inclusion of a new concept word meaningfully related to the inclusive concept category presented during the learning phase. The total design consisted of 8 different treatment conditions.

Subjects

The Ss were 88 undergraduate students enrolled in an introductory educational psychology course at The Pennsylvania State University. Participation in the experiment was voluntary and not a part of the course requirement. The Ss, however, did receive additional credit toward their final grade for serving in the experiment. The Ss were assigned randomly to one of the 8 conditions prior to their arrival at the experimental session. Randomization was recycled at N + 1 treatments.

Lists of Words

The different experimental treatments were attained by variations in the stimulus materials provided in the learning and transfer phases. All lists were comprised of twenty word pairs, made up of combinations of the words shown in Table 1.

In List I of the learning trials the words shown in Column A served as the stimulus elements and those shown in Column B served as the response terms. In List II of the learning trials the position of

Table 1

Words Used in Learning and Transfer Lists

Word-Pairs*		Context Words		
Column A	Column B	Related	Unrelated	Concept-Related
BUILDING	WHISKEY	gin wine	newspaper bus	BEER
CHURCH	MEASLES	polio flu	fruit foot	CANCER
BOOK	ROBBERY	theft assault	pronoun spoon	MURDER
DEN	DRUM	trumpet violin	stove wand	PIANO
PRIEST	WOOD	gas coal	valley spider	OIL
PACK	COPPER	aluminum tin	tornado pear	ZINC
INCH	BANANA	orange peach	well death	APPLE
PAMPHLET	FATHER	brother peach	doctor waltz	AUNT
THUMB	GERMANY	england russia	piano cousin	SPAIN
MAYOR	WOOL	nylon linen	milk boots	SATIN
MILK	DIMES	quarters nickels	hard time	PENNIES
CANOE	TOMATO	pea lettuce	sheet doll	SPINACH
RELIGION	SPARROW	eagle crow	verb bell	ROBIN
FOG	BLOUSE	shoes socks	hermit cough	SHIRT
GIRL	SWORD	bomb club	chemistry key	PISTOL
SAUCE	TEACHER	lawyer dentist	match stamp	SALESMAN
GLACIER	GOLF	tennis swimming	rain mirror	FOOTBALL

(cont'd)

Table 1 (cont'd)

Words Used in Learning and Transfer Lists

Word-Pairs		Context Words		
Column A	Column B	Related	Unrelated	Concept-Related
BIOLOGY	HAMMER	saw nails	meter glass	CHISEL
DOOR	HORSE	lion elephant	sergeant month	DOG
TRUCK	LAMP	bed sofa	emerald boy	DRESSER

* The words in Column A and Column B were used in List I as stimuli and responses, respectively, in each word-pair. The positions were reversed for word-pairs used in List II.

the words was reversed. A context was always present, being conceptually related (CR) in two treatments (List I - CR and List II - CR) or conceptually unrelated (CUR) to Column B words in the other two treatments (List I - CUR and List II - CUR). The context words (listed in Table 1) were typed in small letters and positioned in proximity of the capitalized words shown in Column B as follows:

wine WHISKEY gin	or	newspaper WHISKEY bus

Thus, it can be seen from the above descriptions that the manipulations of function (i.e., stimulus or response) and of context (i.e., related or unrelated) for Column B words were orthogonally crossed resulting in 4 distinct treatments, each represented by a distinct list. Four random orders of presentation were prepared for each list.

The transfer lists consisted of 20 word pairs each consisting of the word from Column A and another word as described below, without the presence of a context. The position (stimulus or response) of words in Column A was interchanged across lists. In the transfer lists, the word embedded within the context during learning was replaced in half the lists by a random selection of one of the original learning trial context elements (CTX); and in the other half the new, concept-related (NCR) words shown in the last column of Table 1. Thus, there were 6 transfer lists with 3 random orders of presentation.

A practice list of 10 paired associates was devised and utilized across all groups. Column A words and practice words were selected at random from any of the 56 categories provided in the Battig and Montague category norms (1969). Corresponding Column B words, related context words, and concept-related words were selected on the basis of high

frequency (1-7) from randomly chosen categories. Unrelated context words were randomly selected from remaining categories.

Procedure

The tasks were administered individually by means of a memory drum. The study-recall procedure was used. The introductory instructions gave S a general orientation to the learning phase of the experiment. All Ss were informed that they would participate in a memory experiment requiring the association of twenty word pairs and oral identification of the second word (response element) when the first (stimulus element) was presented alone; that each testing trial would be preceded by a learning trial exposing both members of the word pair; and that there would be a one-trial practice exercise consisting of 10 paired associates.

The rate of presentation was 3 seconds throughout the experiment. The stimuli were presented until S reached a criterion of one completely correct block of trials. The practice trial was utilized to insure procedural understanding and to reduce possible practice effects in experimental sessions.

After the administration of the practice trial, S was given five minutes to examine the instruction section of Flags: A test of spatial thinking (Thurstone and Jeffrey, 1959). He was then given the second set of instructions which specifically dealt with the presence and function of the accessory words in the forthcoming learning task. Depending upon condition, the context was discussed as proximally related to either stimulus or response elements, but conceptual relatedness (or unrelatedness) of the context to the element of the word-pair was not mentioned. The Ss were informed that, as in the

practice exercise, they would be tested only for verbal recall of the second word in the pair. The S was also told that he could regard the accessory words in any manner he desired. Thus the context could be used as a device to facilitate memory or it could be ignored. The Ss in the four learning groups were then presented a twenty word-pair tape appropriate to the condition to which they were assigned; that is, the S was presented one of the two elements (stimulus or response) in the word-pair. The study-recall trials were terminated at the completion of the 10th presentation or when S reached criterion.

Only Ss who had reached criterion during the learning task were employed in the transfer phase. The unsuccessful Ss were released from the experiment at this time. The transfer instructions indicated that the 20 word pairs would be similar or identical to those formerly experienced in the learning session, but that the context would be eliminated. Depending upon condition, the major element of the word-pair (that is, the element of the word-pair bounded by context words) was replaced by either a context related, context unrelated, or concept-related word. The transfer session concluded upon attainment of the criterion or after ten study-recall presentations.

Results and Discussion

The data were analyzed via a mixed four-factor analysis of variance (2 x 2 x 2 x 2) with three between variables and one within variable. The within variable was the mean number of words recalled by each S in the first trial of the learning phase and the first trial of the transfer phase. The results of this analysis are summarized in Table 2 and the means for the conditions are shown in Table 3. The effects due to trials yielded F (1,80) = 277.52, $p < .001$. The interaction of Learning Context by Trials yielded F (1,80) = 22.38, $p < .001$; and the interaction of Learning Context by Transfer Concept by Trials yielded F (1,80) = 13.32, $p < .001$. The interaction between Learning Context by Position (stimulus or response) by Trials approached significance yielding F (1,80) = 3.62, $.05 < p < .10$. None of the other main effects or interactions were found to be significant ($p > .05$) in this analysis.

The hypothesis that the relatedness of the learning context would produce differential gains for the transfer condition was supported in the Learning Context by Transfer Concept by Trials interaction. Thus, as shown in Figure 1, the related learning contexts though slightly less favorable ($\bar{X} = 7.09$) initially than the unrelated learning context ($\bar{X} = 8.09$), resulted in substantially greater transfer. Though the main effect of Learning Context is not significant during the learning phase, the direction of the data is not consistent with that reported by Pan (1926) who found that initial learning was facilitated by embedding words in a related context. Furthermore, while definite conclusions cannot be drawn from this experiment, the present results

Table 2

Summary of Analysis of Variance of Number of Correct Responses on the First Trial of the Learning and Transfer Phases: Experiment I

Source	df	MS	F
Between Ss			
Learning Context (B)	1	52.36	2.37
Transfer Concept (C)	1	34.57	1.57
Position (S or R) (D)	1	63.84	2.90
B x C	1	12.02	———
B x D	1	0.02	———
C x D	1	0.36	———
Error (b)	80	39.42	
Within Ss			
Trials (A)	1	2385.81	277.52**
A x B	1	192.37	22.38**
A x C	1	5.11	———
A x D	1	2.75	———
A x B x C	1	114.56	13.32**
A x B x D	1	31.11	3.62*
A x C x D	1	26.27	3.05
A x B x C x D	1	23.28	2.71
Error (w)	80	8.60	

** $p < .001$

* $p > .05 < .10$

Table 3

Mean Number of Correct Responses Obtained on the First Learning and Transfer Trials by all Experimental Groups

Groups			Task	
			Learning	Transfer
Context at Stimulus				
Learning context		Transfer word		
Related	-	Context	7.27	17.45
Unrelated	-	Context	7.64	13.55
Related	-	Concept	8.73	16.27
Unrelated	-	Concept	8.64	15.45
Context at Response				
Learning context		Transfer word		
Related	-	Context	5.27	16.55
Unrelated	-	Context	9.09	9.82
Related	-	Concept	7.09	15.91
Unrelated	-	Concept	7.00	14.64

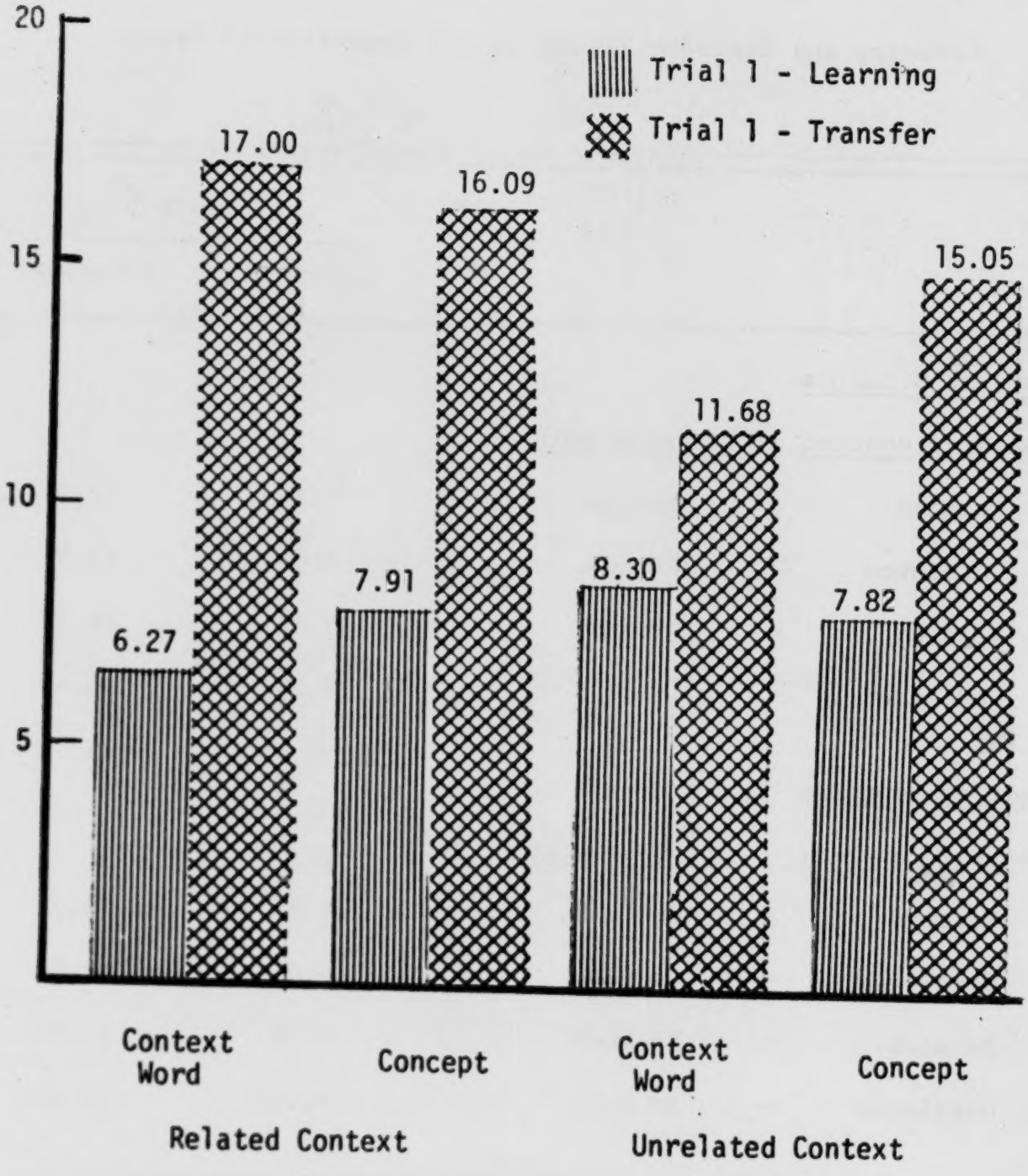


Figure 1. Mean number of correct responses in related and unrelated contexts conditions during the first learning and transfer trials.

suggest that the related learning context tends to induce the learner to conceptualize. This tendency is manifested in the relative inefficient performance in the learning phase. However, in the transfer phase, the earlier conceptualization tends to be facilitative as indicated by the comparatively high scores when Ss responded to specific contextual ($\overline{X}$ = 17.00) and related conceptual ($\overline{X}$ = 16.09) cues.

The Context by Position by Trials interaction, though only approaching significance ($p < .10$), can still be interpreted as non-supportive of the Pan study. The general direction of the data indicate that learning is slightly more favorable when the stimuli are embedded in the context than when the responses are embedded in the context. The trend towards greater increments of improvement in transfer for the stimulus contexts are opposite to the findings obtained by Pan.

The hypothesis that the related context would result in greater transfer to a conceptually related main element than to an unrelated concept was not supported. However, the average transfer scores collapsed across the position variable were slightly higher when the related concept word was learned ($\overline{X}$ = 16.09) than when the conceptually unrelated word was learned ($\overline{X}$ = 15.05).

The results of this study demonstrate the differential effects of context in learning and transfer situations. As a result of this experiment it was reasoned that related contexts are debilitative in the initial learning setting probably because they evoke conceptualizing tendencies on the part of the learner. They, thereby, effect or direct the patterning (encoding) of the learning material by the learner. Since it is the concept that is learned the immediate retrieval of the specific item may be sacrificed to the attainment of the more general

pattern or concept. However, this same tendency in the related context condition appears to be facilitative for later tasks which require a conceptual or categorical placement of the original learning. It apparently makes no difference in this process whether the context is related to the stimulus or to the response element.

Although all of the results were in the predicted direction, some attempt should be made in accounting for the failure to obtain significance for several hypotheses. It is probable that the nature of the paired-associate task is limited in demonstrating the effects of the context variable. The three-second exposure interval may not be optimal for the conceptual processes that are probably elicited by the related context. The obvious categorical relationship of the learning and transfer elements could dilute the effects of context in mediating this relationship. Finally the number of correct responses may not be a sensitive enough measure of transfer. Other measures such as response latency should be investigated in further studies.

Experiment II

This experiment served as a continuation of Part I and was concerned with determining the effects of context in the storage and retrieval of experiences. It was conducted specifically to investigate the effects of context on the conceptualizing tendencies of the S as suggested by an interpretation of the results of Experiment I.

General Design

In a 2 x 4 factorial design with repeated measures one factor was the context during learning; that is whether the context pairs between the stimulus and response elements were related (R) or unrelated (U) to

the response. This factor was crossed orthogonally with four conditions in which the response term, in the transfer task was (a) conceptually-related to both the response element and the context words in the learning task (RCR); (b) conceptually-related to the response element but unrelated to the context words in the learning task (RR); (c) unrelated to the response term in the learning task, although this element was one of the context words shown to the Ss in Group U during the learning trials and had been seen by Group U during the learning trials but was never seen by the Ss in Group R (UCR); (d) a new word which did not appear at any time and was unrelated conceptually to the context or response element during the learning trials (NW). These variations are illustrated in the following chart:

Overall Design

Learning Task	Transfer Task			
	RCR	RR	UCR	NW
Context Related to Response (R)				
[Verb $\begin{matrix}\text{niece}\\\text{uncle}\end{matrix}$ Father]	[Verb-Cousin]	[Verb-Pope]	[Verb-College]	[Verb-Tiger]
Context Unrelated to Response (U)				
[Verb $\begin{matrix}\text{college}\\\text{lunch}\end{matrix}$ Father]	[Verb-Cousin]	[Verb-Pope]	[Verb-College]	[Verb-Tiger]

The repeated measures variable was number of correct responses on the first recall trial in the learning and transfer task in one analysis; and on the first three trials in a related analysis.

Subjects

The Ss were 96 undergraduate educational psychology students. There were 12 Ss assigned to each cell of the design with a constant proportion of males and females (2:3).

General Procedure

All Ss were seated opposite a translucent screen and given the following instructions:

"This is a memory experiment. I (i.e., the Experimenter) will expose on the screen a pair of words written in this position:

MEMORY--JUDGE

Twenty such word-pairs will constitute a series. Your task is to associate each pair of words so as to be able to recall the second word when the first is presented alone. After the series has been presented for the first time, the first word of each pair will be presented alone at the left of the screen and you will be expected to verbally anticipate the corresponding word of the pair. In case you are unable to remember the particular word do not be afraid to guess. At first you may make mistakes, but if you pay close attention you will soon be able to learn which words go together.

When you anticipate a word you are to say it loudly and clearly so I can hear you.

Any questions?

All right, we're ready to begin. Remember the first time through just study the pairs. After that, try to anticipate the corresponding word."

A practice list of 10 paired associates was then presented for one study-recall trial. The E then read Instructions II:

"This phase of the experiment will be the same as the former one except in the following respect. Besides the pair of capitalized words to be memorized, there will be presented two accessory words -- one above and one below the given pair, as shown here:

harm

BALL BENT

take

As in the preceding phase, you will be asked only to anticipate the capitalized word on the right. You may use the accessory words as a memory aid, i.e., to help you remember the pairs or you may choose to ignore them. This is up to you, but remember, you will not be tested for recall of the accessory words in this phase. Remember, too, that the first time through the set just study the pairs.

Any questions?"

A series of 20 word-pairs with either the Related or Unrelated context were then presented for 10 study-recall trials or until S reached the criterion of 20 correct responses in a single trial. The S was then instructed to work on Flags (Thurstone & Jeffrey, 1956) which served as an interpolated activity for a 10 minute interval. The third set of instructions were then given to Ss who successfully reached the criterion. Unsuccessful Ss were released form the experiment. The instructions for this, the transfer phase, were as follows:

"The general procedure in this phase is practically identical to the previous ones. There will be no accessory words given, only the capitalized pair at the left and right sides of the screen. Once again, your task will be to verbally identify the word that appeared on the right. I might point out that this time the pairs will be similar or identical to the ones you studied in the last task.

Any questions?"

The transfer lists were then given with words replacing the response elements. The experiment was terminated when the S completed 8 trials or upon reaching the criterion of 20 correct responses. No S took more than 8 trials. A short interview, consisting of the following questions was administered at the conclusion of the experiment:

1. Did you use any method in particular to help you associate the word-pairs in the first task? (Disregard practice session.)
2. Did learning the words in the first task help or hinder you in the second task? How?
3. Was there anything about the words themselves that helped you to learn the response in the first task?
4. a) Did you notice the accessory words at all while learning? Did you use them in any way?
   b) Did you notice any connection between them and the main elements on the first task?
   c) On the second task?

Results and Discussion

The data were analyzed by a mixed analysis of variance in which the two Learning Contexts (R or U) were crossed with the four transfer conditions (RCR, RR, UCR, NW). The within variable was Trials and consisted of the mean number of words recalled by each S in the first trial or the learning phase and the first trial of the transfer phase.

As shown in Table 4, this analysis yielded F (1,88) = 9.11, $p < .01$, for the effect due to Learning Context; F (3,88) = 7.88, $p < .001$, for the effect due to Transfer Context; and F (1,88) = 193.84, $p < .001$, for the effect due to Trials. The interaction between Transfer Context and Trials yielded F (3,88) = 5.40, $p < 01$. None of the other interactions were found to be significant.

The significant Transfer Context by Trials interaction implies that performance in the transfer task is differentially dependent upon the relationship of the new associate to the original element and its context. Thus, in accordance with the original hypothesis, gains were most favorable for the RCR conditions and least favorable for the NW and UCR conditions. These comparisons are displayed graphically in Figure 2.

In a further analysis of these data the degree of transfer (d) for each experimental group was determined simply by subtracting the learning phase mean from the transfer phase mean. The significance of these differences was then tested by a t test for independent means. In this analysis the NW cell within the U context condition and the NW and UCR cells within the R condition were considered as baseline cells; transfer in these cells was assumed to be zero for these groups, i.e., learning was unaffected by the specific transfer of context or original response relatedness. Strong support for the major hypothesis was demonstrated by the clear superiority of the RCR group (d = 9.66) to the RR group (d = 5.66) in the R context condition (t = 3.28, df = 88, $p < .01$). Of further import was the finding that RCR was greatly superior to NW and UCR ($p < .001$), but RR did not differ significantly

Table 4

Summary of Analysis of Variance of Number of Correct Responses on the First Trial of the Learning and Transfer Phases: Experiment II

Source	df	MS	F
Between Ss			
Learning Context (B)	1	188.01	9.11*
Transfer Context (C)	3	162.57	7.88**
B x C	3	11.24	———
Error (b)	88	20.63	
Within Ss			
Trials (A)	1	1727.99	193.84**
A x B	1	27.00	3.03
A x C	3	48.16	5.40*
A x B x C	3	16.00	1.80
Error (w)	88	8.91	

* $p < .01$

** $p < .001$

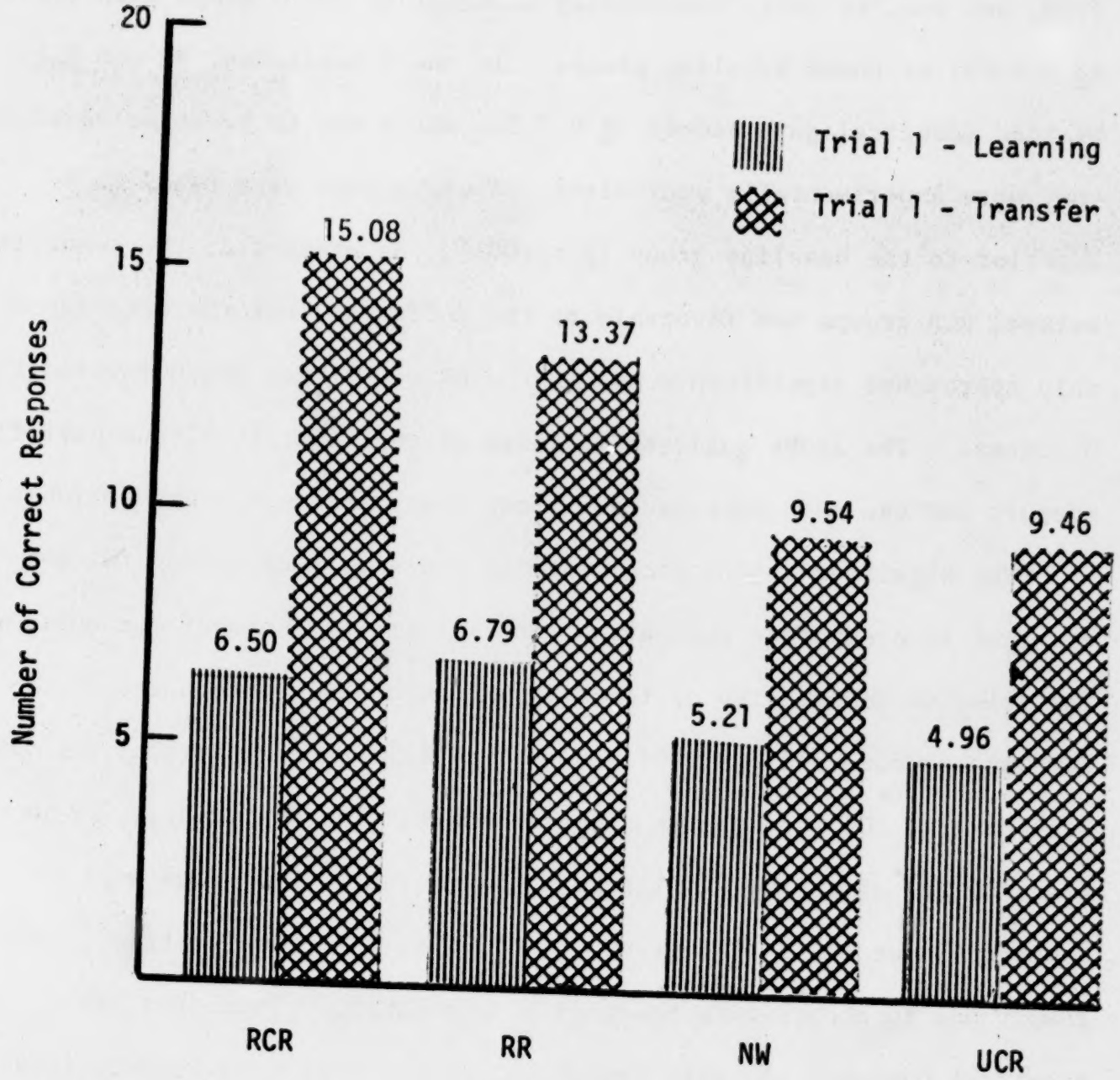


Figure 2. Mean number of correct responses for all transfer conditions.

from, and was, in fact, numerically inferior to the average gain score ($\underline{d}$ = 5.83) of these baseline groups. In the U condition, RR and RCR yielded identical gain scores ($\underline{d}$ = 7.50) which was to be expected since they were experimentally equivalent. These groups were found to be superior to the baseline group ($\underline{p} < .001$). As predicted, the comparison between RCR groups was favorable to the R context, but the difference only approached significance as did the RR comparison which favored the U context. The above analyses indicate strong statistical and directional support for the main hypothesis. Group means are summarized in Table 5.

The significant effect of Learning Context (R or U) was further examined in a separate analysis of the learning and transfer conditions. The relative inferiority of learning the word-pairs within the R context was pronounced ($\underline{p} < .001$) in the initial learning trials, but less extreme ($\underline{p} < .05$) during the transfer phase. The differences during transfer are almost solely attributable to the U (RR) group's clear dominance over the R (RR) group ($\underline{p} < .01$) in the transfer trial. This result was in accord with the initial hypothesis. Thus, for all practical puproses the data imply that both contexts facilitate transfer in equal fashion, despite the relatively poor performance of the R groups during the learning phase. To further investigate this finding, an analysis similar to that described in the immediately preceding paragraphs was performed to detect transfer gain differences between R and U conditions. This comparison yielded a significant effect ($\underline{p} < .05$) favorable to the R condition. The results of the above analyses can be summarized as strongly supportive of the hypothesized differential effects of context in acquisition and transfer. Thus,

Table 5

Mean Number of Correct Responses Within the Learning and Transfer Phases by all Experimental Groups:

Experiment II

Learning Context	Transfer Condition	Learning Phase	Transfer Phase
Related (R)	(RCR)	4.92	14.58
	(RR)	5.75	11.41
	(NW)	3.58	9.00
	(UCR)	3.75	10.00
Unrelated (U)	(RCR)	8.08	15.58
	(RR)	7.83	15.33
	(NW)	6.83	10.08
	(UCR)	6.17	8.91

the related context, though inhibiting performance during the acquisition stages, results in comparable performance to other groups on the transfer tasks.

Discussion

The results of the two experiments reported here imply that specific experiences can, under some circumstances, become embedded within a more general context or cognitive structure. The distinctiveness of the specific experience then tends to be sacrificed in favor of the more general pattern or concept. Thus, a specific item, which stands only as an exemplar of a concept, will not be retrieved (recalled) as efficiently when it is incorporated into a conceptual pattern as it would when merely associated with another item or otherwise stored via Type I transformation (i.e., in more or less arbitrary fashion).* On the other hand, when tasks are performed subsequently which require the recall or application of the concept, substantial transfer can be observed. This implication was only suggested by Experiment I but was strongly supported by the results of Experiment II. The comparatively strong effect of learning contexts in Experiment II may be attributable to the relocation of the accessory words between the stimulus and response terms of the word-pairs rather than above and below one of the terms as it was in Experiment I.

Evidence for the positive effect of the related learning context on transfer was also provided in both experiments. The net gain was

---

* For a description of Type I, II, and III transformations see the article "An Evolving Theory of Instruction" in this Annual Report.

shown to be significantly greater for the related conditions in both studies though the experimental design of the first experiment was less suitable for this type of analysis. Though the related-context groups (i.e., the groups learning word-pairs presented jointly with related accessory words) learned the initial task more slowly than other groups, their performance was equal to that of the unrelated context groups in the transfer phase. These findings suggest a number of other highly interesting questions for further research, such as: Under what conditions can initial learning with contextual frameworks be facilitated? How can the specific item be disembodied from the cognitive structure thereby making it more distinctive and more easily retrieved? What is the extent of the transfer advantage of learning in context as measured by the range of applications that can be made by the learner? Additional studies are certainly required that control the level of initial learning and include, as a baseline, a no-context learning condition.

Perhaps the most interesting of the present results is the strong support of the hypothesis that human learning is a dynamic process ... that learners act on the material to be learned as well as being acted on by the material. Thus, the related context was found to be a powerful determinant of the manner in which the stimulus is coded and stored. Evidence for Type II transformations were clearly found in this study. Thus, in Experiment II it was shown that the learning of a word such as "Father" placed in the context provided by the accessory words ... aunt and uncle ... results in substantial transfer when replaced by "Cousin" but comparatively little transfer when replaced by "Priest." As interesting was the related finding that groups which learned the initial association with an unrelated context could transfer equally

well to either "Cousin" or "Priest." This can be interpreted as strong support for the previously discussed notion that the related context elicits definite "conceptualization" tendencies in the learner. If the unrelated context groups conceptualized to the same extent, their inefficiency in transferring to the "un-conceptualized" meaning (which would have appeared probabilistically fifty percent of the time) would have been evident. It is thus assumed that differences in transfer for related-context groups must result from factors related to a conceptual or categorical structuring of the response during learning. Furthermore, the S does not appear to be aware of these processes; few Ss reported a conscious use or study of either context.

A further question is whether or not the unrelated context is ever incorporated into the learners perception of the stimulus pattern. The above discussion implies that it is rejected or ignored from the outset. The primary evidence for this suggestion comes from the finding that unrelated-context groups are successful in recalling the original element throughout the learning sequence. This finding also implies that the Ss were, for the most part, undistracted. No extra time was required for complex processing. When an unrelated-context word was inserted as a main element during transfer, results from both experiments indicate little recognition of the previously experienced word. In fact, the original context word was not recalled more easily than a completely new word. If the unrelated context is processed in some manner, its association with the stimulus or response element as evidenced in the present transfer task, at least, is negligible. It seems reasonable to assume that the unrelated context is ignored by the learner, or if perceived it is rejected early in learning.

In a more general way, the relationship between the focal stimuli (i.e., the word-pairs to be learned) and the accessory words can be conceptualized as a figure-ground relationship. The focal stimuli are judged by their background, i.e., by their context. Embed the focal stimuli in a related context with characteristics like their own and their distinctiveness is lost; embed it in meaningless context and the background becomes noise, then the figure stands out. It is important to recognize that in the course of these events the meaning of a word becomes subordinated to the context in which it appears; it becomes transformed in the sense that its meaning depends on the context.

These findings and implications are relevant not only to instructional strategies but to study habits as well. What a learner takes down in his note-taking may make a profound difference in what he recalls or in what he can transfer. His notes, in a real sense, betray his transformations. In this regard there will be subtle differences among learners. Some learners will perform acts of omission ... thus, if they jot down only two characteristics of a conceptual context this may not be as precise as three or four, thereby delimiting later ability in the use of the material learned. Other learners will perform acts of elaboration. They will bring their own contexts to the notes they take thereby often modifying the intent of the communication.

The principle to be understood is that context during learning whether provided by the instructor or imposed by the learner can be positively or negatively, or neutrally related to transfer and retention (Heim, 1957; Helson, 1964). As shown in the present study the relationship of the material to be learned to the context can be

an important factor. Fortuitous backgrounds or adventitious contingencies (contexts) create significant transformations on focal items but may themselves be learned incidentally. The first attack on this problem has demonstrated that a context leads to a change in structure and it is the changes structure that is transferred or retained. Thus, what is learned occurs as a result of an elaborate process involving selective attention, pattern matching, and transformations.

In summary, context has been shown to be a significant factor in affecting learning and transfer. The procedures employed in the present experiment appear to be sufficiently sensitive to the effects of this variable as to warrant further experimentation. Other measures of the dependent variable such as latency or response will be investigated to further understand the processes involved since the reaction qualities measured by latency appear to be especially appropriate for the behavioral processes assumed to be employed by Ss in these experiments. Individual differences, especially those involving propensity for conceptualization, should be examined in light of the effects of context for certain types of learners. The present findings strongly suggest that certain groups might benefit differentially from such variations in context as related vs. unrelated, picture vs. word, no context vs. context conditions, as well as kind of context. Later we shall be concerned with (a) influencing designated changes in the focal stimuli by knowing the characteristics of the fringe stimuli; (b) identifying conditions under which concepts (i.e., patterns or codes) contrasted with specific experiences, given the same contextual cues, are retrieved; and (c) examining the effects of differences in rules for storage and

retrieval on transfer. (For the moment the latter may be illustrated by a learning situation in which the material is stored according to conceptual relationships and retrieved according to associational relationships.) These studies can be extended to include such individual differences as the distinctions between "levelers" and "sharpeners" or between "imagers" and "verbalizers."

References

Heim, A. W. Psychological adaptaion as a response to variations in difficulty and intensity. The Journal of General Psychology, 1957, 56, 193-211.

Helson, H. Adaptation - level theory: An experimental and systematic approach to behavior. New York: Harper & Row, 1964.

Pan, S. Influence of context upon learning and recall. Journal of Experimental Psychology, 1926, 9, 468-491.

Summary

The Effects of Presentation Modalities and Modality Preferences on Learning and Recall*

Study Director: Gary M. Ingersoll

Advisor: Francis J. Di Vesta

Technical Problem

This study investigated the conditions under which individuals who differentially prefer to have information presented over one sensory modality as opposed to another, learn and recall stimulus materials presented over the two modalities. The performance of visualizers, i.e., those Ss who preferred to have material presented visually, and listeners, i.e., those Ss who preferred to have material presented auditorily, was compared in a variety of bisensory auditory-visual tasks. It was assumed that in tasks in which unfamiliar materials were presented simultaneously over two sensory modalities, S would be unable to attend to both modalities and therefore he would select one or the other. It was further anticipated that the modality to which he attended was a stable response characteristic. The present investigations were oriented toward the establishment of definable aural and

---

* An earlier progress report entitled "The Effects of Presentation Modalities and Attending Preferences on Learning and Recall" was included in the January, 1970, Semi-Annual Report. The present summary is of a Ph.D. dissertation conducted under this contract. The dissertation will also appear as a Technical Report.

visual modality preferences (during presentation of material by auditory and visual means simultaneously) which are stable across tasks and populations of Ss.

The principal model from which this work was initiated is Broadbent's (1968) limited capacity processing mechanism model. Briefly, the model proposes that a given individual can allow a specific amount of information to enter the processing system within a limited amount of time and that information is processed at a fixed rate. If that rate is superseded by the presentation of information simultaneously across more than one channel, the individual will monitor the flow of information by restricting or closing off the flow from one or more inputs until the rate of input no longer surpasses the capabilities of the mechanism.

An examination of the Broadbent model reveals that little attention is directed toward the role of definable individual differences in the processing of information. Current information processing models, for example the computer paralleling model of Atkinson and Shiffrin (1968) and Shiffrin and Atkinson (1969), offer monitoring systems in which different response strategies or biases are imposed on incoming stimuli. In addressing themselves to the problem of simultaneous inputs, Atkinson and Shiffrin (1968) note, "The first decision the subject must make concerns which sensory register to attend to. Thus, in experiments with simultaneous inputs from several sensory channels the subject can readily report information (from one channel) if so instructed in advance, but his accuracy is greatly reduced if instructions are delayed until after presentation" (Atkinson & Shiffrin, 1968, p. 107). If, however,

no instructions are provided, the individual must impose his own preferences for monitoring information. The extent to which this is done and the stability with which it is done, should be reflected in response output.

General Methodology

The investigation consisted of two independent studies in which visualizers and auralizers were defined on a bisensory auditory-visual task and then compared for performance on additional bisensory tasks. Early studies which have alluded to modality preferences in bisensory presentation, have done so using the first omitted response as their defining response. The first emitted response, albeit a corollary of the original Broadbent (1958) model, is not a sufficiently stable measure under a variety of conditions to warrant its use as the definition of stable individual differences in modality preferences. Senf, Rollins and Madsen (1967), for example, demonstrated that order of response was highly influenced by mental set. Further, early pilot investigations by the present investigator revealed that some Ss develop an effective strategy in which they process the "easiest" modality first and hold it in store while emitting the less preferred modality. The less preferred modality, although emitted first, was not processed first and should still suffer the greatest decay. Thus, if items from one modality were consistently recalled with greater accuracy, that modality was defined as the preferred modality since, by implication, it was the more accurately processed.

Following the definition of modality preferences, visualizers and listeners were selected to participate in a series of test tasks. These

tasks are designed to measure different levels of cognitive functioning. In this way, some evidence were provided which were to describe the pervasiveness of the individual difference in question. The tasks are described, in detail, immediately below.

Missing units task. Two independent sets of five words were presented simultaneously to S, one set was presented visually while the other was presented auditorily. Four words from each set were then repeated on the same modality and S was required to respond with the two missing words, one from each set.

Clustering task. This task was intended to test the strength of the modality preference under the conditions of another well established effect. Six sets of six words which are normatively categorized or grouped were presented to S. During presentation, 18 words were presented on each channel simultaneously, three words from each of the six sets. Following the presentation of the bisensory list, S was given instructions to recall as many of the items in any order that he pleased. Three trials were given.

Paired-associate task. In this task, S was required to learn a list of associates as in a paired-associate task. However, in this case an inter-channel association had to be made. One half of a pair was presented visually; simultaneously, the other half of the pair was presented aurally. An S was required to learn as many pairs and as much of the list as possible within a limited number of trials. A modified study-test procedure was used.

Complex learning task. In this final task, paragraphs of approximately the same length, factual content and familiarity were presented

to S in a bisensory manner. Two independent paragraphs were presented simultaneously, one on each channel for an equal exposure time. The S was then required to recall as many facts as possible from each paragraph. This task defines the maximally dissonant conditions under which modality preference was studied in this investigation and should provide evidence as to the generalizability of the individual difference.

Technical Results

The results show a disordinal interaction between modality preferences and presentation modalities at least in short-term memory. That is, listeners recalled more auditory stimuli than visual stimuli and visualizers recalled more visual stimuli than auditory stimuli. Further, the results offer support for a separate sensory storage model such as that offered by Murdock (1966, 1967). Not only did listeners recall more auditory stimuli, but those auditory stimuli which were presented more recently were recalled better. Conversely, not only did visualizers recall more visual stimuli but those visual stimuli presented in the earlier part of the list were recalled better.

On more complex tasks, the results were not as clearly defined. However, the results of the studies strongly suggest nonlinguistic factors in the effects of modality preference and presentation modalities. With unfamiliar information, a modality and preference interaction was found.

Educational Implications

The present investigation was considered as the initial stage in the development of a theoretical framework within which the generality

and the limits of the construct "modality preference" were to be identified. Eventually, a nomological net, in which this construct is more fully defined should emerge as additional data defining the characteristics of visualizers and auralizers are gathered. Such investigations are indispensable if aural and visual modality preferences, as constructs, are to be incorporated into a theory of instruction, as it eventually must be since so much of present day instructional strategies is dependent upon the presentation of materials via these two modalities.

These data suggest that in settings where information is arriving on more than one channel, individuals differentially sort out or choose one or the other of the modalities and that modality which they choose is a stable characteristic. Thus, we might assume that in areas where audio-visual materials are used in instructional aids and where the information coming over both channels is not entirely congruent (or is somehow different) that some of the information may be lost because of the nature of the multichannel stimulation. This loss may be augmented by the nature of selective attention as employed by each type of individual. Students, therefore, who consistently attend to the visual component of the task will suffer most on demands for information from the auditory channel. Likewise, auralizers may suffer when recall is demanded of visual information. It would appear, then, that in settings of auditory-visual concomitance of presentation more research must be done to delimit these possible effects.

Implications for Further Research

Inasmuch as this was the initial study in the delimitation of modality preferences on bisensory learning and recall, there are many areas that need clarification. Many of the results of this initial set of studies are suggestive and further investigations are warranted. A clearer definition of the role of modality preferences in complex tasks is required. In view of the fact that recency effects were observed for listeners and primacy effects were dominant for visualizers, other investigations are necessary to identify further effects on storage, recall, and retrieval of information.

References

Atkinson, R. C. & Shiffrin, R. M. Human memory: A proposed system and its control processes. In K. W. Spence and J. T. Spence (Eds.) The psychology of learning and motivation: Advances in research and theory. Vol. 2. New York: Academic Press, 1968.

Broadbent, D. E. Perception and communication. New York: Pergammon Press, 1958.

Murdock, B. B., Jr. Visual and auditory stores in short-term memory. Quarterly Journal of Experimental Psychology, 1966, 18, 206-211.

Murdock, B. B., Jr. Auditory and visual stores in short-term memory. Acta Psychologia, 1967, 27, 316-324.

Senf, G. M., Rollins, H.A. & Madsen, M. C. Recall order of dual-channel input as a function of induced recall strategy and stimulus rate. Proceedings of the 75th Annual Convention of the American Psychological Association, 1967, 2, 57-58.

Shiffrin, R. M. & Atkinson, R. C. Storage and retrieval processes in long-term memory. Psychological Review, 1969, 76, 179-193.

Summary

Note-taking and Review in Reception Learning

Donald L. Peters and Carl Harris

Technical Problem

This study investigated the effects of permitting note-taking, distributing prepared notes, or prohibiting note-taking on the learning of technical material from a taped lecture presentation under conditions of review or no review. Much of the previous literature using the reception learning paradigm prohibited such learning relevant activities as note-taking and review and it was hypothesized that such constraints would reduce the amount of learning manifest on a subsequent examination. Possible interactions between such constraints and the individual differences among students were also investigated.

General Methodology

An experiment was conducted where three variations in note-taking and two variations in review time were independently manipulated. Twenty S's were randomly assigned to one of the six treatment conditions and the entire group was administered an individual test battery, a taped lecture presentation, and a posttest on the lecture material. The experimental conditions were manipulated by means of written directions.

Technical Results

A two way analysis of variance (Review by Notes) was performed on the posttest results. The analysis indicated a main effect due to the note-taking conditions but no significant effect for the review conditions. Subjects permitted to take notes during the taped presentation and subjects provided with printed notes performed equally well and both were superior to subjects permitted no notes. Significant main effects on performance were found for the locus of control measure, intolerance for ambiguity, facilitating anxiety and debilitating anxiety. An aptitude X treatment interaction was found for the intolerance for ambiguity individual difference variable and the note-taking conditions.

Educational Implications

The results indicate that the student instrumental activity of note-taking, usually ignored in the reception learning paradigm, is important. Consideration of the activities typically engaged in by the student during the normal classroom situation are necessary for a full understanding of classroom learning.

Implications for Further Research

This study suggests that a more detailed analysis of student note-taking behavior is warranted. Of particular importance would be the investigation of the particular aspects of student note-taking behavior that relate to learning. Such variables as style of notes, quantity of notes, content of notes would seem basic. The relationship of note-taking variables to other individual difference variables should also be investigated.

Note-taking and Review in Reception Learning[1]

Donald L. Peters and Carl Harris

In most classroom situations the material is presented to the student in accordance with the reception learning paradigm (lecture and didactic methods). Yet, little is known about the relationship between this mode of presentation and the learning - relevant instrumental activities engaged in by students. Such behaviors fall into the broad category of behaviors termed mathemagenic behaviors by Rothkopf (1968). This seeming paradox arises, at least partially, because in the typical research situation (see, for example, Ausubel, 1963), the learner is required to process and internalize the material without engaging in many of the standard procedures he would be expected to use in the normal classroom situation. That is, he is permitted neither to take notes for review purposes nor to abstract and organize the material with the aid of written notes. Review time is not usually permitted.

Restrictions upon the usual note-taking and review behavior of students both reduce the generalizability of the results for actual classroom situations and places an unfair and unrealistic burden upon the student. The present study attempts to determine the effects of two typical laboratory restrictions--prohibiting notes and prohibiting review--on the learning of new material from a taped lecture presentation.

---

[1] Appreciation is expressed to the students and staff of the Huntingdon Area High School for their cooperation and assistance in the conduct of this study.

Further, since it is unlikely that the restrictions placed upon student instrumental activities would prove equally deliterious to all students, an analysis of the effects of restrictions in light of individual differences among students was attempted. Interactions were sought between the treatment manipulations and the individual status variables of achievement anxiety, open and close-mindedness, intolerance for ambituity, and locus of control.

Method

Subjects

One hundred and twenty high school juniors served as subjects for the study. They represented the majority of the students in a local high school enrolled in American History. The school serves both low and middle socioeconomic status neighborhoods.

Procedure

The study was conducted in a large room. Initially the Ss were administered a 45 minute test battery which included a pretest on the learning material and a series of aptitude measures. The anonimity of each S was maintained throughout the study by the use of randomly assigned nine digit numbers. The subjects used their numbers instead of their names on all test materials.

The two restriction dimensions were manipulated through a series of written instructions distributed randomly following the initial testing. The general instructions provided for all subjects were:

Your booklet is probably a different color than the one of the person next to you. This means that you have somewhat different directions than the other people in the room. You therefore should follow very carefully the written directions in your booklet, and pay no attention to what the other fellow is doing.

FOLLOW ONLY THE DIRECTIONS APPEARING IN YOUR BOOKLET. THEY ARE SPECIFICALLY FOR YOU.

You are about to hear another tape on some different material, again of college level. (This study was conducted on the same day as another study that also used a taped lecture presentation.) We are interested in how well you can learn this material in a short period of time. Listen carefully to the material as you will be tested on it later.

Now, once again, you each have a set of special instructions to follow. Do not pay attention to what others are doing. Just follow the directions in your booklet. Turn to the next page for your special instructions.

On the subsequent page of the booklet appeared one of the following:

A. Listen carefully to the taped material. Do not take notes.
Do not write anywhere in this booklet. Just listen carefully to the material.
DO NOT TAKE NOTES OR WRITE ANYWHERE IN THIS BOOKLET.
Pay no attention to what others are doing. Just listen carefully to the material.

B. Listen carefully to the taped material. You may take notes on the following yellow sheets. DO NOT WRITE ANYWHERE ELSE IN THIS BOOKLET.
Pay no attention to what others are doing, just listen carefully to the material and take notes.

C. Listen carefully to the taped material. You have provided in the next few pages some notes on the material which you may follow as the lecture proceeds.
DO NOT TAKE NOTES AND DO NOT WRITE ANYWHERE IN THIS BOOKLET.
Pay no attention to what others are doing, just listen carefully to the material and follow the prepared notes.

At this point, a twelve minute tape recording concerning steel alloying (adapted from Ausubel, 1963) was played. During this time the two E's served as proctors insuring, by means of the color coding on the booklets, that the subjects were following their own set of directions.

Immediately following the recording the subjects were verbally directed to proceed to the next page in their booklets where they would find further instructions to follow. The instructions found therein were of two types:

1. Turn to the next page and begin to answer the test questions. Pay no attention to what others are doing.
2. Do not turn to the test that follows until you hear the teacher say "Begin". You may use the interim time to think about the material you have heard (with form A), (or to review your notes (with forms B & C.)

   Remember, pay no attention to what others around you are doing and do not begin the test until you hear the teacher say "Begin."

The command "Begin" was given 4½ minutes after the end of the taped session. The test instructions prohibited the subjects from turning back in the test booklet. After all subjects had completed the test, they were collected and the entire group was returned to their normal classroom routine.

Measures

The pretest consisted of 15 five-alternative multiple choice questions relating to the taped material. The fact that this material was entirely new to the subjects is supported by the near chance level of responses found on this measure. The mean number of items correct was 4.3, and the internal consistency reliability ($r = .09$) did not differ significantly from zero.

The criterion measure was a 25 item, five-alternative multiple choise test. The internal consistency of this measure was found to be .42 ($p < .05$). The majority of the items on both tests were adapted from Ausubel, (1963).

The aptitude test battery consisted of four measures which yielded five scores. The measures were: 1) the Internal-External Scale (Rotter, 1966); 2) the Achievement Anxiety Scale (Alpert & Haber, 1960) which yields scores for facilitating and debilitating anxiety; 3) the Dogmatism Scale, Form E (Rokeach, 1960); and 4) the Intolerance for Ambiguity Scale (Budner, 1963).

Results

The intercorrelation of the measures used in the study are presented in Table 1. It can be seen that small but significant correlations were found among several of the aptitude measures and between the locus of control, intolerance for ambiguity, and achievement anxiety measures and the posttest scores. No significant correlations were found with the pretest scores.

A two way analysis of variance (Review X Note Conditions) was performed on the posttest results. The analysis indicated a main effect due to the note-taking conditions but no significant effect for the review conditions. The interaction between the two restriction dimensions was also non-significant. Three way analyses of variance (Aptitude X Review X Note Conditions) were also performed and indicated no significant interactions between the review condition and any of the aptitude variables. Therefore, for purposes of clarity, only the two way analyses (Note Conditions X Aptitudes) are reported here. The subjects within each of the note conditions were dichotomized at the median value of the aptitude variables for each analysis.

Table 2 reports the results of the analysis for the dichotomized locus of control measure and the note conditions. It can be seen that

Table 1

Intercorrelation of the Measures
(N = 120)

Measures	2	3	4	5	6	7
1. Locus of Control	.07	.11	-.23**	.20*	-.10	-.17*
2. Intolerance for Ambiguity		.19*	-.21*	.17*	-.16	-.23**
3. Dogmatism			.11	.31**	.01	-.15
4. Facilitating Anxiety				-.35**	.15	.14
5. Debilitating Anxiety					-.16	-.35**
6. Pretest						.13
7. Postests						

* $p < .05$

** $p < .01$

Table 2

Analysis of Variance for
Locus of Control X Three Notes Conditions

Aptitude	No-Notes		Notes		Prepared Notes	
	$\bar{X}$	s	$\bar{X}$	s	$\bar{X}$	s
Internal	7.95	2.52	9.15	3.60	8.85	2.25
External	6.40	2.33	8.10	2.65	7.80	2.73

Effect	df	Mean Squares	F Ratio
Notes	2	23.43	3.17*
Locus of Control	1	44.41	5.99**
Interaction	2	0.83	0.11
Error	114	7.40	

* $p < .05$

** $p < .01$

the aptitude and treatment variables both yielded main effects significant beyond the .05 level. The aptitude X treatment interaction was not found to be significant.

The results indicate that the main effect of the note conditions may be attributed to the significantly poorer performance of the subjects who were permitted no notes. No significant difference exists between the Notes Permitted and Prepared Notes conditions. Those subjects scoring low on the aptitude measure (internal locus of control) performed better than those scoring high on the measure in all treatment conditions.

Table 3 indicates the results of a similar analysis, this time dichotomizing the subjects on the basis of their intolerance for ambiguity scores. The results indicate that among subjects scoring low (tolerant) on the intolerance for ambiguity measure performance without notes was inferior to performance in the other two treatment conditions. However, among persons scoring high on intolerance for ambituity, there were no significant differences in performance in the three conditions of note-taking. This interaction is depicted in Figure 1.

Tables 4 and 5 present the analyses for the dichotomized achievement anxiety scores. The results indicate the expected main effects due to anxiety with subjects rated high on facilitating anxiety scoring better than subjects rated low on this measure and subjects rated high on debilitating anxiety scoring significantly lower on the criterion measure than subjects with low debilitating anxiety scores. No significant interactions were found.

A similar analysis was performed on the dichotomized dogmatism scores but the results indicated that neither the main effect of the dogmatism variable, nor the interaction were significant at the .05 level.

Table 3

Analysis of Variance for
Intolerance for Ambiguity and Three Note Conditions

Aptitude	No-Notes		Notes		Prepared Notes	
	$\bar{X}$	s	$\bar{X}$	s	$\bar{X}$	s
Tolerant	6.60	2.30	9.70	3.20	9.40	2.72
Intolerant	7.75	2.65	7.55	2.78	7.25	1.80

Effect	df	Mean Squares	F Ratio
Notes	2	23.43	3.41*
Intolerance	1	33.08	4.81*
Interaction	2	36.30	5.28**
Error	114	6.88	

* $p < .05$

** $p < .01$

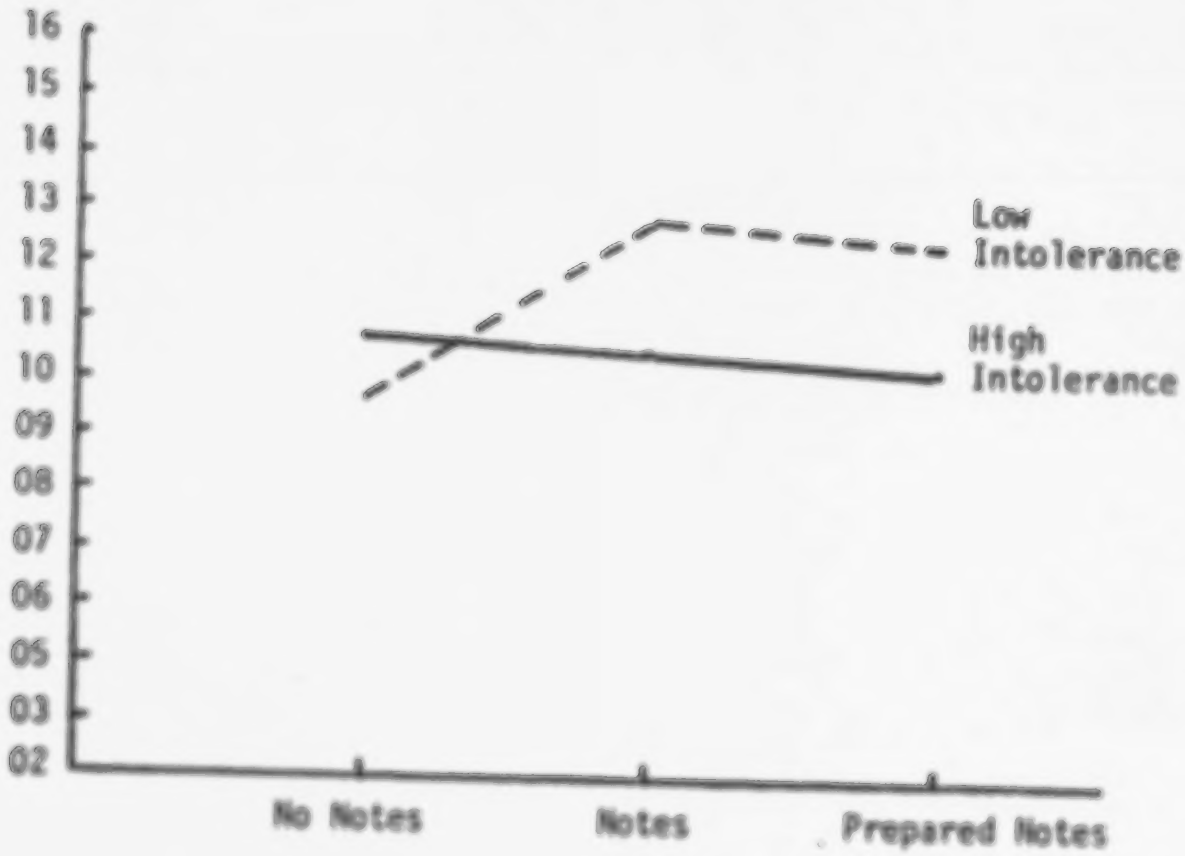


Figure 1. Relationship of aptitude to outcome measures across three note-taking conditions.

Table 4
Analysis of Variance for Dichotomized
Facilitating Anxiety X Three Note Conditions

Aptitude	No-Notes		Notes		Prepared Notes	
	$\bar{X}$	s	$\bar{X}$	s	$\bar{X}$	s
Low Fac. Anxiety	6.65	2.54	7.95	2.50	8.00	2.51
High Fac. Anxiety	7.70	2.45	9.30	3.62	8.65	2.56

Effect	df	Mean Squares	F Ratio
Notes	2	23.43	3.12*
Fac. Anxiety	1	31.01	4.18*
Interaction	2	1.23	0.16
Error	114	7.51	

* $p < .05$

Table 5

Analysis of Variance for Dichotomized
Debilitating Anxiety X Three Note Taking Conditions

Aptitude	No-Notes		Notes		Prepared Notes	
	$\bar{X}$	s	$\bar{X}$	s	$\bar{X}$	s
Low Deb. Anxiety	7.40	2.64	9.90	3.31	8.90	2.36
High Deb. Anxiety	6.95	2.44	7.35	2.50	7.75	2.61

Effect	df	Mean Squares	F Ratio
Notes	1	23.43	3.30*
Deb. Anxiety	2	57.41	8.08**
Interaction	2	11.43	1.61
Error	114	7.10	

* $p < .05$

** $p < .01$

Discussion

Two sets of restrictions typically imposed upon subjects during studies using the reception learning paradigm were studied to determine the effects of such constraints of the normal task-relevant mathemagenic behaviors of students. The results indicate that prohibiting students from taking notes during a taped lecture significantly interferes with their performance on a subsequent recognition test. Subjects in the No-Notes condition were found to score at or very near the chance level on a 25 question five-alternative multiple-choice examination administered either immediately or shortly after the presentation of the material.

The deleterious effect of prohibiting the taking of notes cannot be attributed directly to the effort of note-taking itself, or to more careful attention paid to the presentation during the act of taking notes. Providing the students with a brief outline of the material which they could follow during the lecture served as well as permitting them to actively engage in note-taking behavior. No significant differences were found between the Prepared Notes and Note-taking conditions.

The lack of significant effect of the review conditions could be attributed to either the short duration of the review time provided or the lack of motivation of the students to fully cooperate. The brief review time (4½ minutes) would have permitted only a cursory review of the notes the subjects had available. This may have prohibited any systematic attempt to adequately study the material. Further, there was no opportunity to insure that the students who were provided review time actually used the time to review. While precautions were taken to see that subjects did not proceed with the test until told to do so, no assurance can be given that they used their time productively.

The analyses of the effects of the aptitude variables indicate the relevance of these to performance in a reception learning situation. Four of the five aptitude (student status) variables produced significant effects on the learning outcome.

Subjects with an internalized locus of control outperformed their external locus classmates. The data are consistent with the notion that the internal locus person performs more diligently than the external locus person when there is no external compulsion or reinforcement for doing well. In the present situation the anonimity of the subject's responses and the obvious unrelatedness of the study to regular school work removed most of the externally imposed incentives for achievement. Under such conditions the external locus of control subjects learned very little.

As would be predicted from the theory underlying the Achievement Anxiety Scale (Alpert & Haber, 1960) debilitating anxiety was found to be negatively related to performance while facilitating anxiety was found to be positively associated with performance in the learning situation.

The interaction of the intolerance for ambiguity scores with the treatment variables arose from the poor performance of the low scorers when not permitted to take notes. This result was counter to expectations. In the No-Note condition subjects were directed not to take notes, and yet they undoubtedly were aware that some persons in the room were taking notes or shuffling through papers. It was assumed that subjects in this predicament would define their situation as ambiguous. Therefore, it was predicted that subjects highly intolerant of ambiguity would suffer a decrement in performance under such circumstances. The opposite was found.

An alternative interpretation of the situation is based upon the assumption that the person who is intolerant of ambiguity will, when forced into an amibguous situation, reduce ambiguity through whatever means are open to him. In the experimental situation the best course of action open to such a person may have been to follow his own direction exactly, ignoring all others, and to listen very carefully to the material.

Feather (1969), Crandall (1969) and Peters and Messier (elsewhere in this report) have also reported difficulties in interpretation of the Intolerance for Ambiguity measure. The growing number of unanticipated results suggests that the construct measured by Budner's scale is in need of further definition and possibly a new label.

In conclusion, it appears that the student instrumental activities that are usually ignored or prohibited in the typical reception learning paradigm are important. Their consideration is necessary to a full understanding of the learning processes involved in classroom learning. The results suggest that further investigation of student note-taking behavior is warranted.

References

Alpert, R. and Haber, R. Anxiety in academic achievement situation. Journal of Abnormal and Social Psychology, 1960, 61, 207-215.

Ausubel, D. Cognitive structure and the facilitation of meaningful verbal learning. Journal of Teacher Education, 1963, 14, 217-221.

Budner, S. Tolerance of ambiguity as a personality variable. Journal of Personality, 1963, 30, 29-50.

Crandall, J. Self-perception and interpersonal attraction as related to tolerance-intolerance of ambiguity. Journal of Personality, 1969, 37, 127.

Feather, N. Attitude and selective recall. Journal of Personality and Social Psychology, 1969, 12, 310.

Peters, D. & Messier, V. The effects of written reinforcement and question sequence upon objective test performance. ARPA project report 1970, included in this report.

Rokeach, M. The open and closed mind. New York: Busic Books, 1960.

Rothkopf, E. Two scientific appraoches to the management of instruction. In R. Gagne and W. Gephart (Eds.), Learning research and school subjects. Itasca, Illinois: Peacock, 1968.

Rotter, J. Generalized expectancies for internal versus external control of reinforcement. Psychological Monographs, 1966, 80, whole No. 609.

Summary

The Effects of Written Reinforcement and Question Sequence Upon Objective Test Performance

Donald L. Peters and Victor Messier

Technical Problem

This study involved the investigation of two sets of variables, question sequence and written reinforcement, upon the objective test performance of students. Of particular concern was whether prior experience with tests of a particular construction sequence or with a particular form of written teacher reinforcement would affect performance on subsequent tests. Additionally, concern was directed to finding out if such experimental manipulations would affect individuals with varied personality characteristics differentially.

General Methodology

Two versions of a test (one paralleling the lecture sequence and one containing a random ordering of the same questions) were administered to the subjects on four occasions. Following the scoring of the first three the subjects were cycled through three forms of written reinforcement: no comment, standardized comment, or personalized comment. The next test in the sequence served as the dependent variable for analysis of the effects of the experimental manipulations. The subjects were also administered a battery of individual difference measures.

Technical Results

The results indicate that the particular form of the test had little effect on performance at the time. However, where subjects had prior experience with a random form they did equally well on the next test whether it was random or sequentially ordered. When subjects had prior experience with the sequentially ordered version they suffered a decrement in performance when presented next with a randomly ordered form. No main effects were uncovered for the written reinforcement variable. Several aptitude by treatment interactions were noted.

Educational Implications

The results imply that the question order of a test may have subtle effects on the subsequent study behavior of students. They also suggest that building tests to follow the order of presentation of the material may reduce the generalizability of the learning. The interaction of the individual difference variables with both the test form and the written reinforcement provided by the instructor re-emphasizes the importance of adapting instructional techniques, including the fairly subtle one of this study, to the individual characteristics of students if maximal performance is to be attained.

Implications for Further Research

This was an exploratory study and caution must be exercised in the interpretation of the results. However, it does suggest that investigation of the effects of tests, in terms of both the expectations created and the reinforcement provided, could be profitably studied, particularly in relation to the effects of such manipulations on the subsequent study habits or information processing behaviors of students.

The Effects of Written Reinforcement and Question Sequence Upon Objective Test Performance

Donald L. Peters and Victor Messier

One of the concerns of teachers when constructing objective classroom tests is, "How shall the items be ordered?" A typical answer provided by standard texts on the subject is, "The order in which test items are arranged in the final form of the test is not critical." (Ebel, 1965 p. 157) It has been suggested that early items be made less difficult in an attempt to alleviate anxiety or that the order of items follow the structure of the material as presented by the teacher. When the latter possibility is followed there is the possibility of inter-item cuing or that performance might be enhanced through organizational facilitation of memory and recognition of material in context. To confirm these assumptions, it was hypothesized that students taking a standard, sequentially ordered test form would score significantly higher than those students taking the randomized test form, wherein the items appeared in a randomized fashion relative to the order of class presentation of the material.

The sequence in which course content is presented frequently reflects an externally imposed (though not arbitrary) organization of the material. This extrinsic organization has pedagogical purposes, accommodating the constraints of the instructional system, but it is of secondary importance to the utilization of the learned material. Rothkopf (1968) has proposed that task relevant behaviors such as effort, inspection, search, and review, by engaging the student in active processing of information, facilitate acquisition and achievement. For the learner task irrelevant behaviors,

such as focusing on the sequence of presentation, rather than on the inherent organization of the material, may hinder or potentially hinder learning.

Frase (1968a,b,c) indicated that the events that occur just before the learner is exposed to new material can strongly influence what is learned. He found that pre-test questions act as directive influences and help to establish cues for the learner in deciding what is to be learned. These lead to increased vigilance on the part of the learner and a modification of study habits.

This suggests that test questions which parallel the pedagogical ordering of the material may serve to focus attention on the sequencing of the material and, may be dysfunctional for the learning of new material where performance is measured in situations where the extrinsic ordering is not maintained. Conversely, testing sequences which deemphasize the extrinsic organization may serve to make the learner more vigilant towards the intrinsic organization of the new material, or force him to more actively process the material to provide his own organization. This activity should facilitate recall or recognition of the information in subsequent testing situations no matter what organization the later tests represent. Thus, it was hypothesized that (1) students having been previously tested on items following the sequence of instruction (henceforth such a test will be called "standard form") score significantly lower when later tested using a random test form, i.e., one where the order of items is randomized, than when again administered a standard test form, and (2) students previously tested using a random test form do not score differently on either a randomized or standard form of a subsequent test.

Reinforcement has long been established as a means for increasing task persistence and effort. The motivating value of written teacher comments on objective tests has been established by Page (1958). Such reinforcement, in conjunction with knowledge of results (the corrected tests) is seen as: 1) increasing the student's efforts in preparation for subsequent tests, and 2) providing information as to the relevant content and level of abstraction to be studied. The reinforcement previously awarded on returned tests should have an effect on subsequent test performance. Therefore, a third hypothesis, then, is that students receiving written comments on their returned tests would score significantly higher on a subsequent test than those students who received no comments. Finally, it was hypothesized that the personalized written comments would be more effective than the standardized written comments.

This would suggest that experience with tests which do not follow the extrinsic organization of the course content, coupled with reinforcement in the form of teacher comments, should maximize performance. However, recent research (for example, Alpert & Haber, 1960; Ehrlich & Lee, 1969; Fillenbaum & Jackman, 1961) indicate that not all students will respond equally well to reinforcement or to attempts to break away from the instructionally provided organizational set. Individual differences in tolerance for ambiguity, dogmatism, locus of control, and test anxiety level are likely to influence the acceptance and effect of such procedures. The major hypotheses, therefore, must be examined in light of these individual difference variables.

Relevance For Instruction

As with all instructional procedures, objective tests should be designed and used to accomplish specified educational objectives. They provide a framework for motivating students to search, study, and review essential material and to actively process the material in ways that will provide the greatest long range benefit. The present study should provide evidence of the efficacy of two easy and direct procedures by which an instructor may manipulate the task-relevant and task-irrelevant mathemagenic behaviors of students.

Methods

Subjects

The subjects were 41 graduate students enrolled in a basic research methods course. Twelve were males. They represented a variety of academic fields, with the majority pursuing studies in Child Development and Family Relations. All were enrolled in a beginning course in research methodology.

Measures

A 76 item, four-alternative multiple choice pretest was given to assess the initial level of competence of the subjects. (4 students with scores above the 90th Percentile were eliminated on this basis.)

Three twenty item, four-alternative multiple choice quizzes were administered during the course and a sixty-item multiple choice test, covering the entire content of the course, served as the final examination.

All pretest items, and the majority of the quiz items had been previously used and were found to be good items in terms of discrimination

and difficulty level. Internal consistency reliabilities ranged between .52 and .76.

In addition, as part of the introduction to the measurement section of the course, a short test battery was administered. This included: 1) the Internal-External Scale (Rotter, 1966); the Achievement Anxiety Scale (Alpert and Haber, 1960), which yields scores for facilitating and debilitating anxiety; the Dogmatism Scale, Form E (Rokeach, 1960); and the Intolerance for Ambiguity Scale (Budner, 1962). The scores for each student were recorded and the tests were discussed during subsequent class periods.

Procedure

Each of the three quizzes and the final examination were organized in two ways. The same questions appeared on both versions. One form was organized so that the questions followed the sequential ordering of the course material as it was presented in class. The other form was a randomized version of the first. The two forms were randomly distributed to the students at the beginning of each test period.

After quiz 1 had been marked and grades assigned to the T-score distribution, all the students in a particular grade category (A, B, or C) were randomly assigned to one of three groups; no comment (NC), standard comment (SC), and personal comment (PC) - (after Page, 1958). The NC group had their papers returned with no comments. The SC group had their papers returned with the following standardized comments for the grade levels of A, B, and C, respectively: A - "Excellent! Keep it up"; B - "Perhaps try to do still better!"; and C - "Let's raise this grade!"[1] The PC group had their papers returned with a personal

1. Page had standardized these comments for those students in his study who had achieved the grades A, C, and F, respectively.

comment from the instructor. The comments were addressed to the student by name, referred to personal information about the student's past educational history, and were signed by the instructor. The same procedure was followed for quizzes 2 and 3, with the subjects on the three original groups sequenced through each of the reinforcement conditions. For example, students who received no comment on Quiz 1, received a standard comment on Quiz 2 and a personal comment on Quiz 3.

At the end of the course, the students were questioned to determine if they had been aware of the experimental manipulations. It was apparent that none were aware of the actual intent of the study. Each student was given a mimeographed statement outlining the rationale and design of the study.

Results

Treatment Effects

Because of the small sample used in this preliminary study, and a chance irregularity in the distribution of subjects to treatments by random assignment, it was necessary to collapse cells across the major treatment dimensions for analysis. Therefore, the main effects of the reinforcement and test form dimensions of the analysis were handled separately. This precluded the possibility of investigating any interactions between the two major treatment dimensions.

It was hypothesized that the sequentially ordered versions of the tests would prove less difficult than the randomized versions in all cases. This was tested by one-tailed t tests. Only the results for Quiz 2 were found to be consistent with the hypothesis ($t = 1.89$, $df = 39$, $p < .05$). The results for Quizzes 1 and 3 were in the same direction, but did not reach an acceptable level of significance. The means and standard deviations for each testing are presented in Table 1.

Table 1

Analysis of Effects of the Two Test Forms

Form		Quiz 1	Quiz 2	Quiz 3	Final
Standard	$\bar{X}$	15.52	15.95	14.85	42.91
	s	2.27	1.94	2.48	5.74
	n	20	23	20	21
Random	$\bar{X}$	14.80	14.50	14.33	43.20
	s	2.40	2.88	2.06	5.64
	n	21	18	21	20

It was also hypothesized that the form taken on a prior test would ultimately affect the performance on a subsequent test through modification of the student's study habits. Testing experiences which deemphasize the extrinsic, teacher provided, organization of the material (randomized form) were thought to facilitate recall of the information in subsequent testing situations no matter what organization they represented. Hence, it was predicted that students having a random form on the prior test would do equally well on either form in a subsequent test situation. Students having the standard form on the prior test were predicted to perform less well when this was followed by a random form test than when followed by another standard form test.

To test these hypotheses, the sample was divided into four groups for each pair of testings. That is, they were divided as to the form of the prior test and the subsequent test. The results of the analysis are presented in Table 2.

It can be seen from Table 2 that students who had the sequential form of the test showed differential performance on the two forms during the next testing. By contrast, the students taking the randomized form were not differentiated in their performance on the two test forms on the subsequent testing. In the cases of Quiz 2 and 3, the results obtained were as predicted, with persons having a sequential prior test performing less well on a subsequent random form. In the case of the final examination, the results indicate superior performance on the random form of the subsequent test.

It was hypothesized that the written reinforcement given the student when his tests were returned would have a motivating effect leading to subsequent improved performance on the next testing. Further, it was hypothesized that the personalized comments would be more

Table 2

Effects of Prior Test Form on Subsequent Test Performance

Quiz 1 Form	Quiz 2 Form				
	Sequential		Random		
	$\bar{X}$	s	$\bar{X}$	s	t
Standard	16.53	1.60	14.29	2.00	2.58*
					df=18
Random	15.20	2.00	14.8	2.66	n.s.

Quiz 2 Form	Quiz 3 Form				
	Sequential		Random		
	$\bar{X}$	s	$\bar{X}$	s	t
Standard	16.33	1.60	14.40	1.82	2.25*
					df=21
Random	14.36	1.82	15.33	2.50	n.s.

Quiz 3 Form	Final Examination Form				
	Sequential		Random		
	$\bar{X}$	s	$\bar{X}$	s	t
Standard	42.80	4.40	49.50	4.33	2.80*
					df=19
Random	42.91	5.66	41.98	4.91	n.s.

*$p < .05$

effective than the standardized, and that both forms of reinforcement would be more effective than no reinforcement. To test this hypothesis, a series of one-way analyses of variance were conducted on the criterion measures. The results appear in Table 3.

Page's (1958) findings were not replicated. There were no significant differences found among the reinforcement treatment conditions.

It was suggested that individuals with different aptitudes would respond differentially to the treatments imposed. The distribution of aptitude variables and their relations to the criteria are presented in Table 3. Significant relationships were found between at least one of the criteria and dogmatism, debilitating and facilitating anxiety, and pretest scores.

To analyze the interaction of the aptitude measures and the treatment effects, the linear regression of each of the criteria measures on each of the aptitude measures was computed separately by treatment. Comparison was then made to determine if the slopes of the regression lines varied significantly by treatment for each aptitude and criterion pair. Seven of these were found to be or approach significance. The four interactions involving the test forms and the three interactions involving the reinforcement conditions are summarized in Tables 5 and 6, respectively.

Figure 1 graphically displays the first interaction summarized in Table 5. The interaction suggests that persons scoring high on debilitating anxiety perform less well on tests that have the questions in random order than they do on tests that follow the sequential order of instruction. The opposite is suggested for people scoring very low on debilitating anxiety.

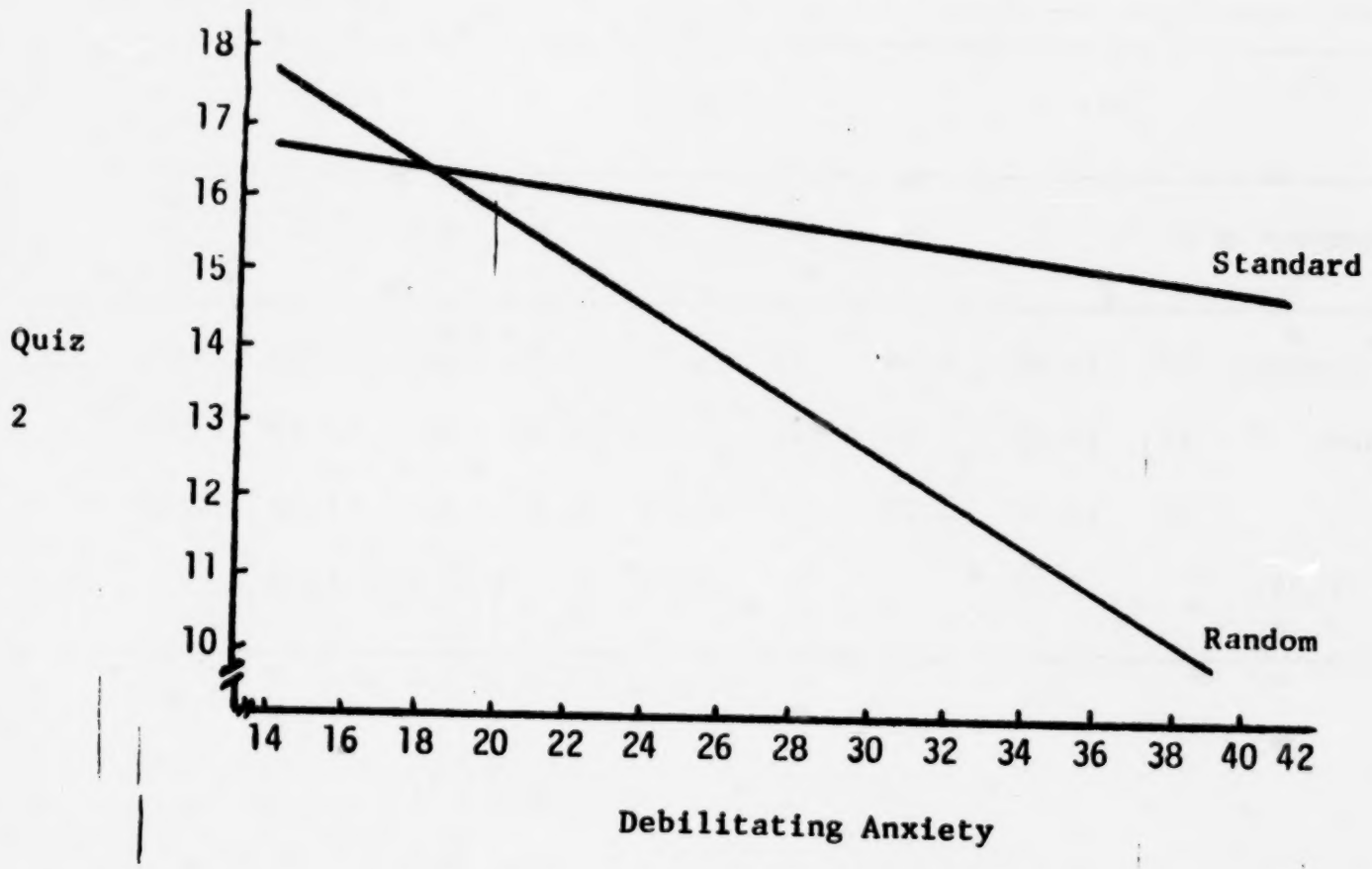


Figure 1

Regression of Quiz Two Scores on Debilitating Anxiety

Table 3

Analysis of the Effects of Three Types of Written Reinforcement on Classroom Test Scores

	Quiz 2			Quiz 3			Final		
Treatment	N	$\bar{X}$	s	N	$\bar{X}$	s	N	$\bar{X}$	s
No comment	17	14.88	2.64	12	15.75	2.30	12	42.58	6.37
Stand.	12	14.68	2.18	17	14.88	2.52	12	45.50	5.05
Pers.	12	16.58	2.27	12	15.00	2.17	17	41.65	5.18
F. (2.35)		2.38			0.52			1.77	

Table 4

Distribution of Aptitude Variables and Their Correlation with the Criteria: Total Sample (N = 41)

Variable	$\bar{X}$	s	Quiz 1	Quiz 2	Quiz 3	Final
Locus of Cont.	9.56	5.19	-.06	-.03	-.02	-.04
Tol. Ambiguity	50.32	7.63	-.05	-.03	-.10	-.03
Dog.	129.83	22.64	-.11	-.18	-.28	-.25
Deb. Anx.	26.32	6.14	.04	.40**	-.26	-.35**
Fac. Anx.	24.95	4.65	.04	.33**	.16	.23
Pretest	39.98	8.56	.30*	.58**	.42**	.53**

*p < .05
**p < .01

Table 5

Regression Equations of Major Interactions

with Test Form

TEST FORM	Criteria	Aptitude	Intercept	Reg. Coef.	df	F
Sequential	Q2	Deb. Anx.	18.36	-.09		
Random			21.98	-.29	1.37	2.90*
Sequential	Q2	Pretest	13.24	.07		
Random			2.29	.30	1.37	14.84***
Sequential	Final	TOA	62.68	-.40	1.37	8.89***
Random			29.91	.26		
Sequential	Final	Dog.	58.03	-.11	1.37	4.44**
Random			35.55	.06		

*p < .10
**p < .05
***p < .01

It should be noted that the sample mean on debilitating anxiety was 26.32 and, although the axes of Figure 1 represent the actual range of scores obtained, only two scores actually fell below the intersection of the two regression lines. This suggests that if an educational decision concerning the administration of tests to students was to be made, the best choice would be to administer the sequential form of the test to all students.

A highly significant interaction between the regression lines for the two treatments was found when the Quiz 2 scores were regressed on the pretest scores. (Figure 2) Subjects performing low on the pretest performed better on the sequential version of the test than on the random, and subjects scoring high on the pretest scored better on the random form than on the sequential form.

Figure 3 represents the regression of the final examination scores on the scores for intolerance for ambiguity. This is a clearly disordinal interaction with the intersection of the two regression lines occuring at approximately the mean of both variables. This interaction indicates that those subjects who were intolerant of ambiguity performed better on the random version of the test, while those scoring low performed better on the sequential form of the test. This interaction is counter to the direction expected.

Inspection of Figure 4 again reveals a disordinal interaction that is counter to expectations. It may be interpreted as indicating that those who score high on the dogmatism scale perform better on the random form of the test than they did on the sequential form. Conversely, those scoring low on the dogmatism scale performed better on the sequential form than they did on the random form.

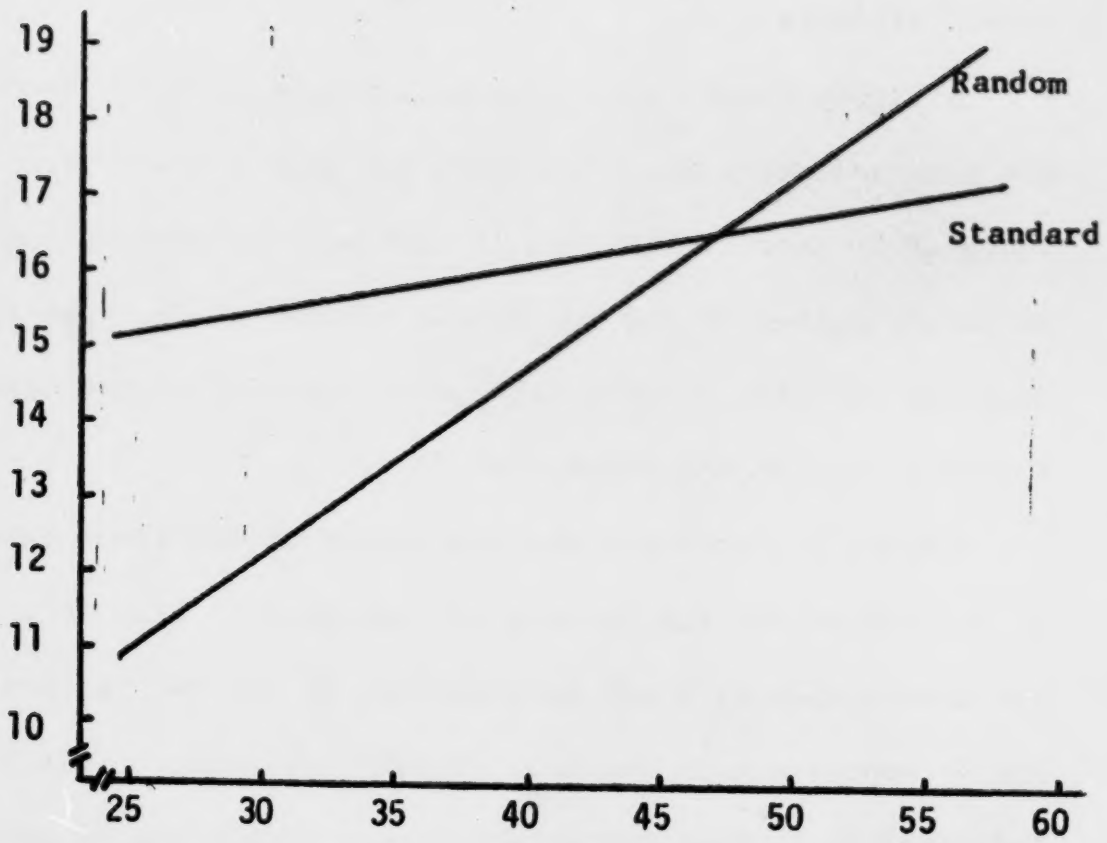


Figure 2

Regression of Quiz 2 Scores on Pretest Scores

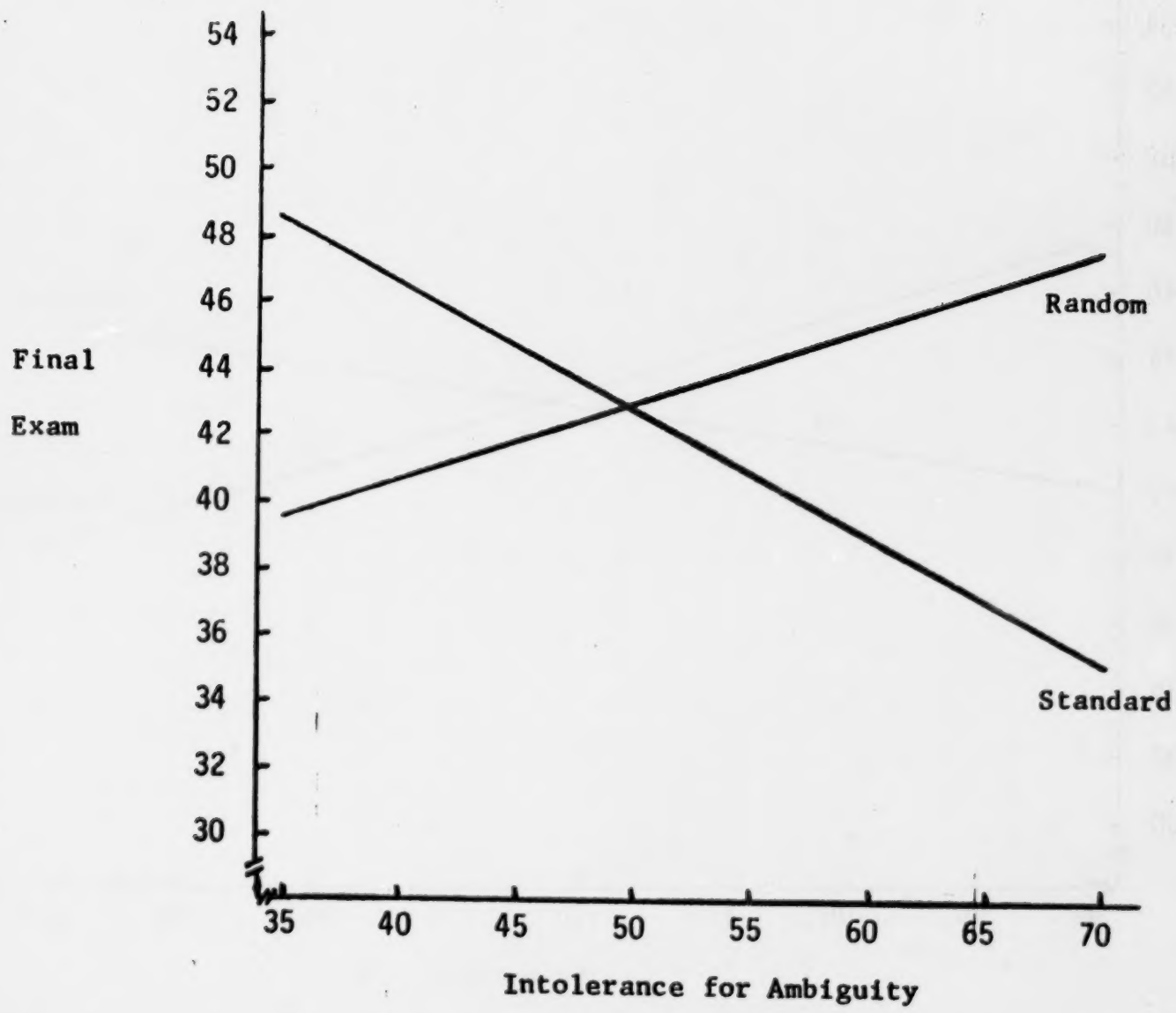


Figure 3

Regression of Final Exam Scores on Intolerance for Ambiguity

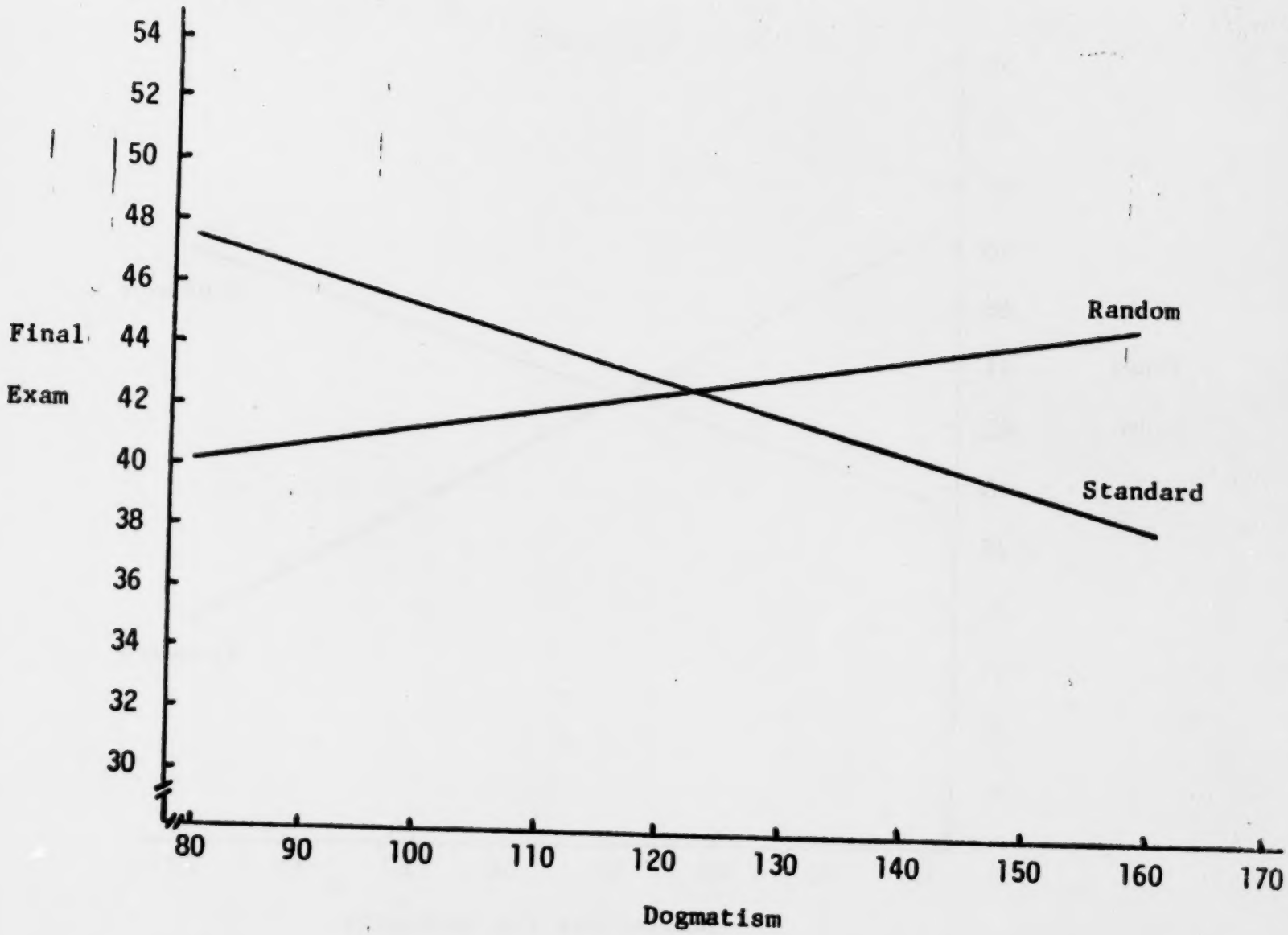


Figure 4

Regression of final examination score on Dogmatism

Table 6

Regression Equations of Major interactions with

Reinforcement Conditions

Treatment	Criteria	Aptitude	Intercept	Reg. Coef.	df	F
No Comment	Q2	LOC	16.63	-.20		
Stan. Comment			13.07	.16		
Pers. Comment			15.98	.06	2.35	2.40*
No Comment	Final	LOC	36.16	.64		
Stan. Comment			46.92	-.14		
Pers. Comment			45.91	-.49	2.35	4.38**
No Comment	Final	TOA	89.00	-.99		
Stan. Comment			35.67	.19		
Pers. Comment			39.29	.05	2.35	7.13***

*$p < .10$
**$p < .05$
***$p < .01$

Figure 5 represents the first three regression lines described in Table 6. There is no difference between the slopes of the two written reinforcement treatment lines. It appears that persons scoring high (External) on the locus of control measure benefit from written reinforcement more than do persons scoring low (Internal) on the locus of control measure.

The analysis of the regression of the final exam scores on the locus of control scores yields exactly the opposite picture. As can be seen in Figure 6, a high score on the locus of control measure is associated with better performance in the No Comment condition. The interaction appears to be disordinal.

The last interaction involves tolerance for ambiguity. As may be seen in Figure 7 and Table 6, this interaction suggests that the higher the score on the tolerance for ambiguity measure, the greater the differential effect of the three reinforcement conditions.

Discussion

The results of this study indicate that if the constructor of a classroom test is concerned only with whether ordering test items randomly or parallel to instruction makes a difference, the most reasonable answer is, "No, it does not." Only one of the four comparisons testing this effect was significant at the .05 level.

However, the overall results contradict the conclusion that the order in which test items are arranged is not critical. Such a conclusion is an oversimplification which does account for the possible utility of the test for directing subsequent learning. Further, it does not take into account the possibility that the arrangement of the questions on a test may be critical for some individuals.

The data presented in Table 2 strongly suggest that the form of the test that the person experiences, in an actual classroom situation, affects in some way his subsequent learning of new material and/or his performance on subsequent tests. Exactly how test experience affects performance is not clear from this research. However, the data are consistent with the notion that the test serves to increase the task-relevant information processing of the student or to inhibit task-irrelevant processing, or both. Frase (1968) has suggested that prequestions serve as cues to identify relevant content for study and that this is accompanied by inhibition of responses to incidental stimuli. More recently, Frase, Patrick, & Schumer (1970) have reported that the negative consequences of such stimuli selection (resulting from prequestions) can be stronger than the direct instructive effects of prequestions. Generalizing this conclusion to the present data would suggest that experience with the randomized versions of the test would decrease attention paid to the extrinsic structure of the material during subsequent learning and, hence, reduce the detrimental effect on performance demands which do not follow the instructional order.

The data do suggest that the use of randomized tests yields learning which has greater generalizability to subsequent occasions. Reliance upon an externally imposed structure of the material may be unrealistic in terms of the subsequent demands made for the knowledge in applied situations. Certainly, procedures which enhance the recall of knowledge outside the instructional context would be consistent with the basic goals of education.

There are a number of possible reasons why Page's finding of the significant effects of written reinforcement were not replicated in this study (Table 3). Page used junior and senior high school students,

whereas beginning graduate students were the subjects in this study. The motivation of beginning graduate students taking their first course already may be at such a high level that reinforcement, particularly in the form of written comments, would not be strong enough to produce any differential effects.

Table 4 indicates a clear relationship between debilitating test anxiety and test performance for three of the four criteria. The greater the person's debilitating anxiety, the poorer his performance. The relationship is especially marked for the final examination and the second quiz. Since the final examination weighed heavily in determining the course grades of the students, this result is entirely consistent with the theory behind the Achievement Anxiety Scale.

Quiz 2 involved the use and understanding of statistical concepts, and the significant negative correlation between student's scores on it and their debilitating anxiety scores is consistent with both theory and intuitive impressions as to what is anxiety producing for graduate students. The interaction displayed in Figure 1 further clarifies this relation. When achievement anxiety is high, and when the situation is presumably anxiety producing, performance on a randomly ordered test is inferior to performance on a sequentially ordered series of the same items. The higher one's achievement anxiety, the more detrimental is the effect of departure from the externally imposed order of the material.

Figure 2 suggests that reliance on the extrinsically imposed order of the material is greatest for those subjects who initially knew least about the subject matter. When this order was destroyed, their performance was markedly inferior. While this may indicate that sequential instruction is an important aid for the more naive student, it also

indicates the fragile nature of their grasp of the knowledge. It would certainly be a mistake to assume that the performance of such students on a sequentially ordered version of the test reflects a firm grasp of the subject matter.

It was assumed at the outset of this research that the random form of the tests represented an ambiguous situation for students. That is, lack of structure was equated with ambiguity. As such, it was thought that subjects with a high intolerance for ambiguity would perform more efficiently on the sequential (structured) versions of the tests. The results indicate just the opposite occurred. It is difficult to interpret why this should have happened, but one very tentative explanation is offered. If the initial assumption was correct, then it is possible that the subjects ranking high on this variable may have responded to the ambiguity by working with increased diligence and care in an attempt to reduce the ambiguity. It would have been informative to have recorded the length of time taken by each subject to complete the test. This might have provided the additional clues required to support this interpretation.

A central position of Rokeach's (1960) theory of the organization of belief-disbelief systems is that the cognitive system of closed-minded persons (high dogmatic) is highly resistant to change. Further, a difference between high and low dogmatic persons is hypothesized in their dependence upon authority. Open-minded persons should be more able to distinguish the source of information from the quality of information than is the close-minded person. By a somewhat extensive extrapolation from the original theory, it was expected that the open-minded subject would be better able to differentiate the extrinsic structure of the material from the content itself and, hence, he should

be less affected by attempts to move away from this structure.

Again, the results obtained were contrary to the prediction. One not altogether convincing possibility is that the open-minded persons differentiated the structure from the content more than did the closed-minded persons but, at the same time, they accepted the structure as reasonable and, hence, tied their learning firmly to it. They would, therefore, be at a disadvantage when the structural support for their learning was removed. At the same time, other research suggests that the closed-minded person paradoxically spends more time studying incongruent materials (Smith, 1968). As with persons rated intolerant of ambiguity, it may be that the high dogmatic persons took more time with the random form of the test and, hence, performed more adequately.

The first interaction among the reinforcement conditions was noted when the Quiz 2 scores were regressed on the Locus of Control scores. One would expect the written comments to have little effect on those students with an internal locus of control, but a reinforcing effect on those students with an external locus of control. And subjects with an internal LOC should score higher than subjects with an external LOC under the no comment (control) condition. The interactions shown in Figure 5 support the above statements.

However, when the final exam scores are regressed on the Locus of Control scores, the results are just the opposite of what was expected (see Figure 6) and what was supported with Quiz 2. Again, it is difficult to find a plausible interpretation for such contradictory results. But one explanation may be that the students who received no comments on Quiz 3 had received personal comments on Quiz 1, and conversely with the subjects who received personal comments on Quiz 3, thus indicating that the initial reinforcement conditions had long term effects which

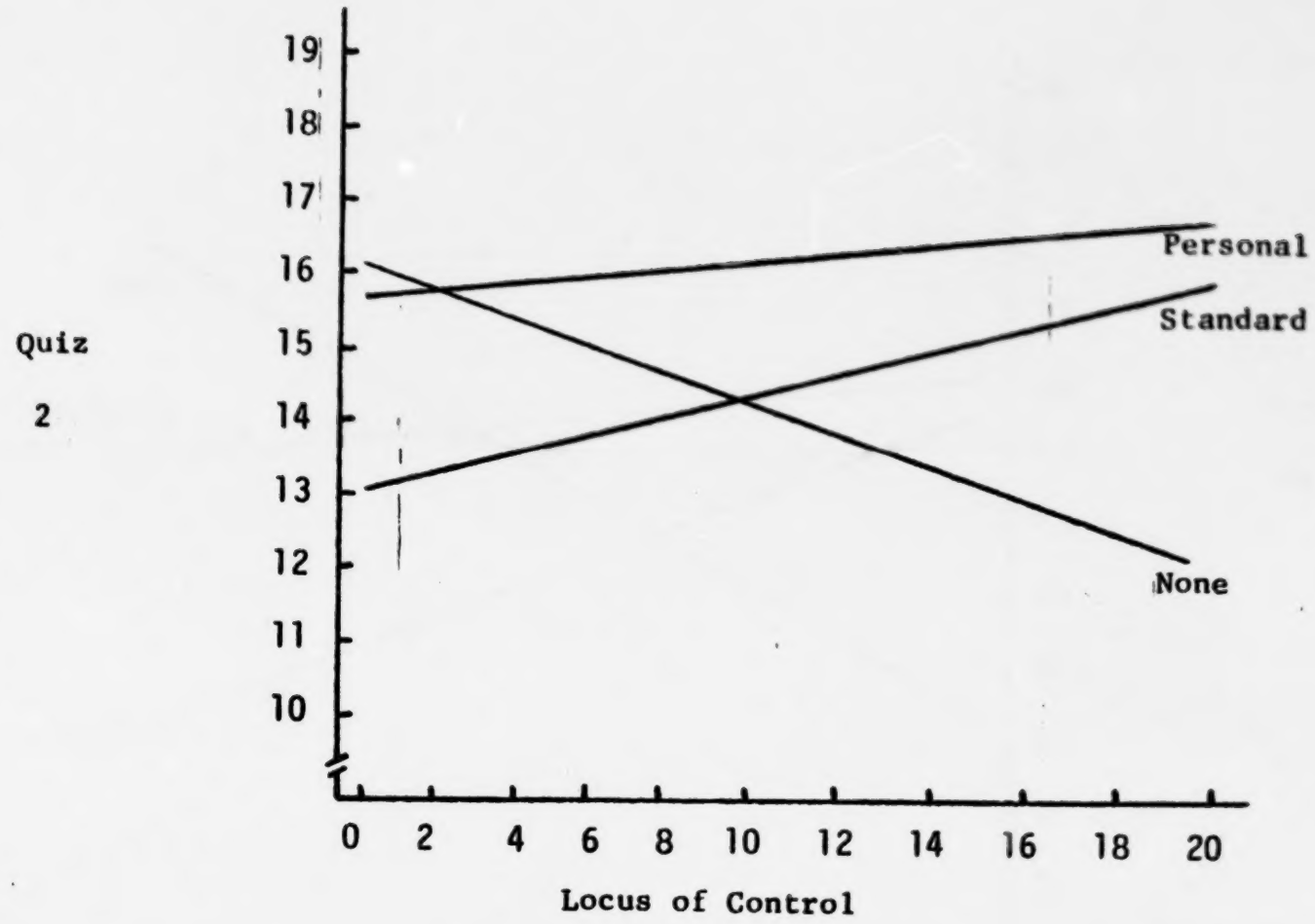


Figure 5

Regression of Quiz 2 Scores on Locus of Control

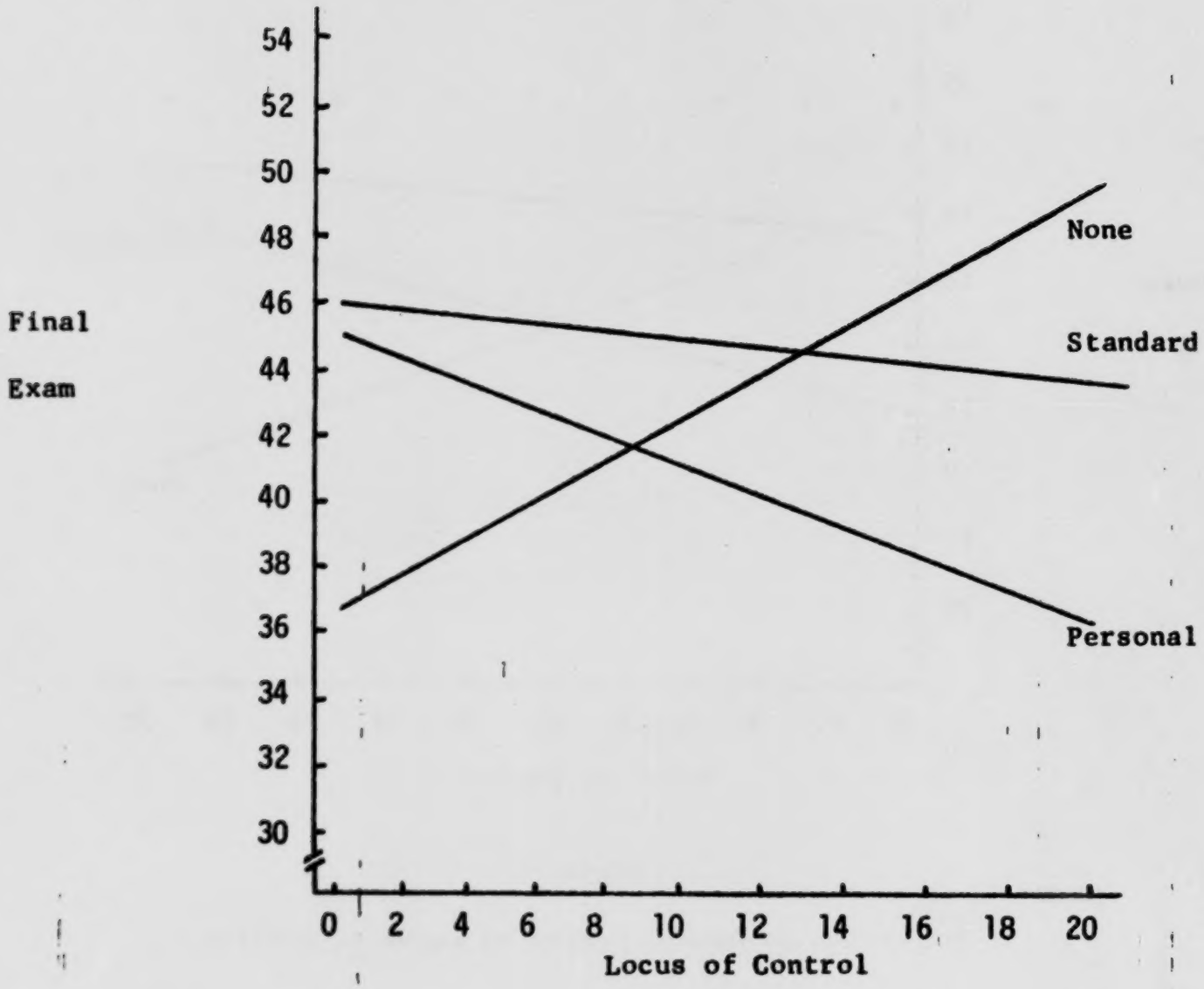


Figure 6

Regression of Final Scores on Locus of Control

became evident when the testing situation was of sufficient importance, as in the final exam. That is, there may be a primacy effect which overrides later attempts to manipulate the reinforcement variable.

The final interaction, and perhaps the most difficult to explain is shown in Figure 7. One explanation for the low scores of the subjects who received no comment and who have high intolerance for ambiguity is that these subjects had received comments on each of the preceding quizzes, and when they received no comments on Quiz 3 their debilitating anxiety increased, resulting in lower performance.

It is clear that there is nothing in the data or the theory which can lead to an assured interpretation of these results. The resolution of these difficulties will have to await further research.

Two remaining points are worth noting. Firstly, in the interpretation of the results it should be remembered that this study involved multiple administration of the treatments. On each occasion the tests were distributed randomly, and by the final examination the great majority of the subjects had experienced both randomized and sequential test forms. Although the data in Table 2 suggest that the immediately previous test had an effect on the next test results, the combined effects of multiple treatments could not be assessed.

Secondly, the correlations in Table 3 and the fact that significant interactions were only found with Quiz 2 and the final examination scores as criteria suggest the importance of motivation in obtaining results in such studies. Had the study been conducted under conditions with less ego involvement for the subjects, the results might have been lost. The recent study by Frase et al. (1970) has indicated the importance of incentives in determining the effects of questioning upon learning. It is apparent that one asset of studies of the present

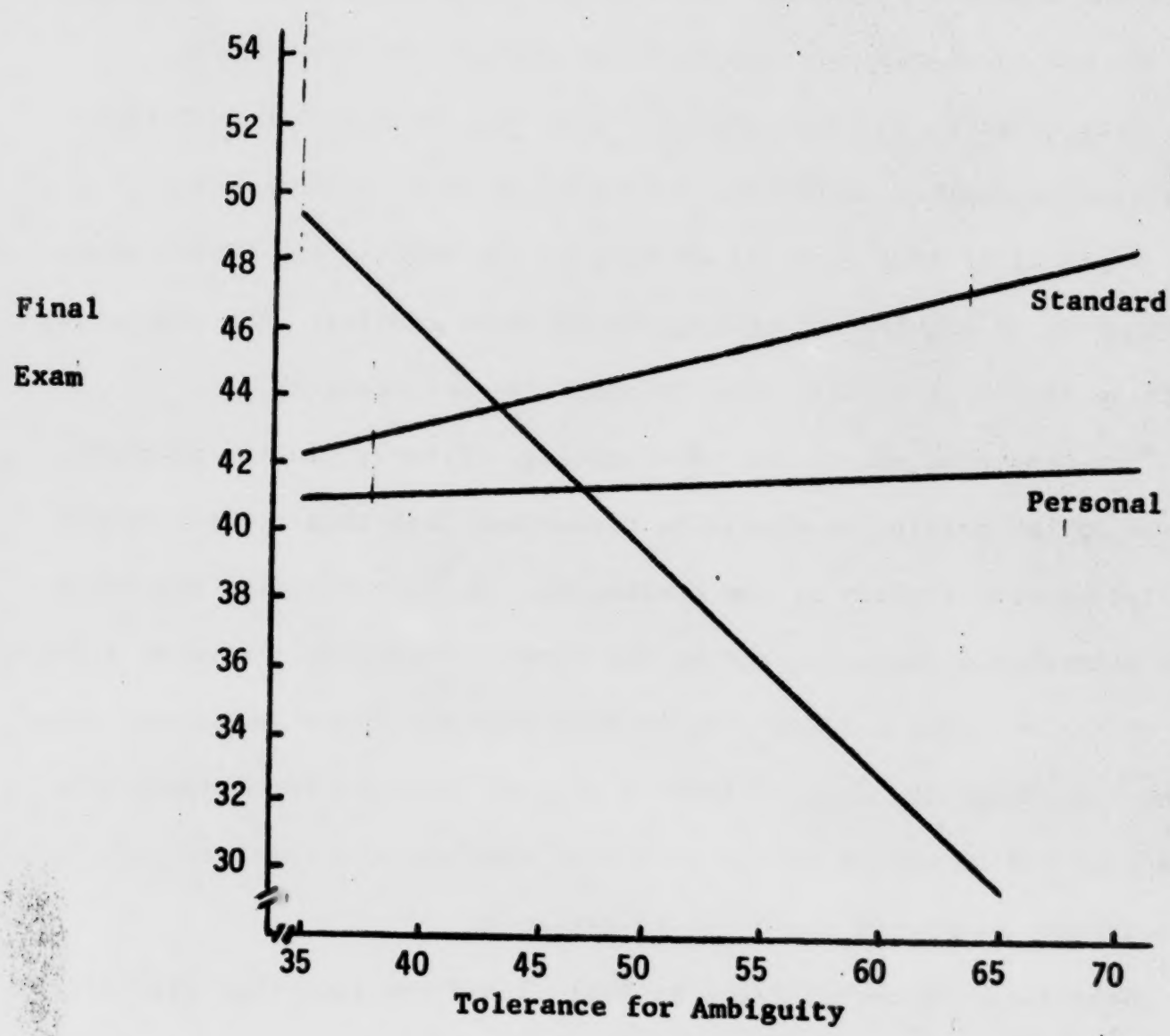


Figure 7

Regression of Final Scores on Tolerance for Ambiguity

type, conducted as part of the on-going classroom experience of students, is the seriousness and high level of motivation of the subject. It is under such conditions that individual differences that relate to learning are most likely to play a role.

Summary

The order of items on a series of four classroom multiple choice tests was manipulated to provide two forms. One form corresponded to the order of classroom presentation of the material, the second represented a randomization of the same items. Little evidence was found to support any overall differences between the two forms. However, experience with one or the other form was found to significantly effect performance in subsequent testing situations. Reinforcement in the form of written teacher comments did not result in significant changes in test performance. The results were explained in terms of the mathemagenic behaviors of students, particularly the inhibiting effect of the randomized version on the learning of the extrinsic structure of the material. Individual differences between students were studied as possible factors effecting responses to the two test forms. Interactions were found between the treatments for the regression of the criteria scores on debilitating anxiety, pretest, intolerance for ambiguity and dogmatism.

References

Alpert, R. and Haber, R. Anxiety in academic achievement situations. Journal of Abnormal and Social Psychology, 1960, 61, 207-215.

Budner, S. Tolerance of ambiguity as a personality variable. Journal of Personality, 1963, 30, 29-50.

Ebel, R. Measuring educational achievement. Englewood, N. J.: Prentice-Hall, 1965.

Ehrlich, H. and Lee, D. Dogmatism, learning, and resistance to change. Psychological Bulletin, 1969, 71, 249-260.

Fillenbaum, S. and Jackman, A. Dogmatism and anxiety in relation to problem solving. Journal of Abnormal and Social Psychology. 1961, 63, 212-214.

Frase, L. Effect of question location, pacing, and mode upon retention of prose material. Journal of Educational Psychology, 1968, 59, 244-249 (a).

Frase, L. Questions as aids to reading: some research and theory. American Educational Research Journal, 1968, 5, 319-332 (b).

Frase, L. Some unpredicted effects of different questions upon learning from connected discourse. Journal of Educational Psychology, 1968, 59, 197-201 (c).

Frase, L., Patrick, E. and Schumer, H. Effect of question position and frequency upon learning from text under different levels of incentive. Journal of Educational Psychology, 1970, 61, 52-56.

Rokeach, M. The open and closed mind. New York: Basic Books, 1960.

Rothkoph, E. Two scientific approaches to the management of instruction. In R. Gagne and W. Gephart (Eds.), Learning research and school subjects. Itasca, Illinois: Peacock, 1968.

Rotter, J. Generalized expectancies for internal versus external control of reinforcement. Psychological Monographs, 1966, 80, whole No. 609.

Summary

The Effects of Studying Together and Grading Procedures on Recall of Subject Matter

Nicholas M. Sanders

Technical Problem

The relative efficiencies of studying together and studying alone were compared under three performance evaluation conditions: direct competition with a study partner (competitive), combining the learner's score with that of a study partner (cooperative), and general competition with all persons taking the performance measure (normative). The degree of dogmatism and the degree of achievement anxiety of the learner were predicted to be important individual differences in assessing the effectiveness of studying together.

General Methodology

The six treatment conditions defined by the two study procedures and three grading procedures were manipulated experimentally in a small group laboratory setting. After an introduction to the experiment which included the instruction as to grading procedure, Ss heard an eighteen minute, taped lecture on details of the lives of three psychologists. A study session of ten minutes followed, in which Ss in the study together condition were allowed to exchange information by writing on 5 x 8 cards. Then a 60 item, short answer test was administered with a fifteen minute time limit.

Individual differences in dogmatism and anxiety were measured in a large group testing session six to eight weeks prior to the time of

participation in the experiment. However, the E had no knowledge of Ss' scores until after the experimental session was completed.

Technical Results

Neither manipulation of study procedure nor manipulation of grading procedure produced overall effects, though the interaction between the variables was significant. Direct competition resulted in the highest mean amount of subject matter recalled under study together conditions and the second lowest mean score under study alone conditions. In the cooperative grading procedure the study together Ss produced the lowest mean score and the study alone Ss had the second highest mean score. There was no difference between the study conditions for Ss in the normative grading condition. These results are contrary to predictions, which were based on research indicating contrasting effects of cooperation and competition on pooling of information. An interpretation of the results obtained is that the grading conditions affect the learner's perception of his chances of success, which are also influenced by the opportunity to gain information about the capability of his "classmate" in the study together condition.

The degrees of dogmatism and anxiety of learners did not differentially relate to individual performance in the treatment conditions. The lack of such relationships was attributed to their conceptual irrelevance to the new interpretation of the effects of the treatment variables.

Educational Implications

The post hoc interpretation of results given above implies that students will do better than usual if they are to be graded by combining

their score with that of another unknown classmate, while students will perform more poorly than usual when they are to be graded by comparing their score with that of another unknown classmate. However, if given the opportunity to study with that other classmate, the students in the two grading conditions will reverse their relative positions. When the student is graded in comparison with all others in the class, the effect of studying with another student does not exist.

These implications are contingent on results of future research, which is described below.

Implications for Further Research

Several aspects of the study require further elaboration. The interaction revealed was explained after the fact, and the explanation requires additional empirical support. The additional research should include modifications of the present study to allow for assessments of perceived chances of success during the progress of the experimental session. In addition, measures of need for achievement and the achievement anxiety variable should be used in combination to assess the individual differences appropriate to the new interpretation.

If the additional research results in support for the interpretation of present study, the task and study period should be modified to determine whether the interpretation is generalizable to modifications in task and study period.

The Effects of Studying Together and Grading Procedures on Recall of Subject Matter

Nicholas M. Sanders[1]

One type of student activity that might be instrumental to learning subject matter is studying together with other students. Many instructional, personality, and cognitive variables undoubtedly must be considered in order to arrive at a detailed statement of the conditions under which studying together facilitates learning. The present research is an investigation of the effectiveness of studying together as a function of one instructional variable, grading procedure, and three personality variables, debilitating anxiety, facilitating anxiety, and dogmatism.

The social psychological research on group versus individual problem solving provides a framework in which to consider initially some of the variables that are important in determining the relative effects of studying together and studying alone. At least two features of the group setting emerge as potentially important. First, there is the possibility for a pooling of information and interpretations in the group setting. And, second, the social aspect of the setting may lead to increased arousal, leading either to greater persistence in work on the task or to disruption of work on the task.

---

[1]The author gratefully acknowledges the assistance of Mr. Sam Rock in collection of the data.

When students prepare for an examination together, there would seem to be a distinct advantage over studying alone since the students may add to one another's knowledge by providing information not noted or remembered by the others and by checking one another's interpretations and applications of the information. However, the "pooling" of information advantage may be mitigated by the evaluation procedure used by the instructor, whose grading system often places students in competition with one another. Research in group problem solving suggests that the procedure for reporting quality of task performance analogous to grading procedures has an effect on the extent to which information is offered and trusted. Deutsch (1949) conducted an experiment in which introductory psychology course discussion sections of five students each were told either that their section as a whole would receive the same grade, based on overall quality of work on problems posed (cooperative condition) or that each student in the section would receive a different grade, based on a comparison of his contributions with those of the other four students in his section (competitive condition). Ratings by observers of the discussions indicated that while there was more said in the competitive condition, there were more frequent misunderstandings and requests for repetitions of what was said, indicating a greater lack of attentiveness in the competitive condition; ratings by the students themselves were in agreement with the observers' ratings. More ideas were agreed upon and adopted as a basis for further discussion and action in the cooperative group. However, no significant differences were found on indices of individual performance, as revealed in observers' ratings of the quality of the student's contribution to the discussions and in grades on assignments done outside of class.

While Deutsch's study does reveal that a cooperative orientation induced by reporting quality of performance in terms of overall group performance does lead to greater communication among group members, it does not allow for a clear test of increased personal performance as a direct effect of the communications. Zander and Wolfe (1964) used a set of problems that required utilization of information obtained from others in the group, and therefore provides more definitive results in terms of effects of reporting procedures on information utilization. They manipulated the reporting procedure by posting for their employee subjects' bosses, either the employees' individual score, the score of the employee's group, or both the individual and the group scores. In agreement with Deutsch's findings, they reported more information relevant interchanges and more trust in others expressed in the group score and combination score conditions, indicating that the group interchange potential was utilized more effectively in those conditions than in the individual score conditions. On the criterion measure of information utilization, which was more directly related to the group interchange than was Deutsch's criteria, Zander and Wolfe found that highest personal scores were obtained by the combination score subjects, while the individual score condition resulted in the lowest personal scores.

In comparing the results of the two studies, one must compare the degrees of cooperation and competition within the group. At the competitive extreme is Deutsch's competition condition: group members are in direct competition with others <u>in</u> the group. The other extreme is exemplified by Deutsch's same grade condition and Zander and Wolfe's group score condition. Zander and Wolfe's individual score condition is in between the two extremes. Though the subject was in competition

with others in all groups, including his own, he might have chosen to be cooperative with those in his group in order better to compete with those in other groups. (The normative grading procedure used in formal education settings is most similar to Zander and Wolfe's individual score procedure: while the student is in competition with all others in the class, he may choose to work with other classmates to improve his chances of success over the remainder of the class.) Zander and Wolfe's combination condition is not on the cooperative-competitive continuum, since that condition would allow, without conflict, both cooperation and competition.

Using the above comparison of conditions on a cooperative-competitive continuum, one can see that Deutsch found no differences in information utilization between the two extremes, while Zander and Wolfe found differences between an intervening condition - the individual score condition - and a condition representing the cooperative extreme. In view of the greater group interchange relevance of Zander and Wolfe's criterion, the present author concludes that a cooperative grading procedure should lead to a higher level of individual performance than would a competitive grading procedure under conditions allowing for group exchange of information.

While the above discussion has dealt with the effects of manipulating motivation to provide and accept information, it is probable that there are individual differences that also affect provision and acceptance of task-relevant information. The dogmatic individual, who evaluates information primarily on the basis of the source of the information (Rokeach, 1960; Powell, 1962; and Ehrlich and Lee, 1969) should tend not to trust information supplied by his peers, and, therefore, would not

profit from group settings, and his relative lack of benefit should become greater in situations where pooling is very advantageous, i.e., cooperative settings.

In addition to the possibilities for pooling of information, the group task setting may have a motivational characteristic that differs from the individual task setting. Zajonc (1965) proposes an arousal level theory to resolve seeming contradictions in research comparisons of production of correct response under group and individual conditions. Zajonc views the presence of others as a source of arousal. If the correct response is readily available to the individual, then the presence of others, by increasing the individual's arousal level will increase the probability of that response occurring. If however, the most available responses are not correct, the probability of the correct response occurring is less in the presence of others than in the individual setting. Assuming that when students study together for a test and the correct responses are for the most part not readily available, the arousal effects of the social setting would result in overall debilitation of performance.

Some individuals, however, seem less debilitated than others by arousal related to academic achievement. Alpert and Haber (1960) propose that achievement-related anxiety may not only be differentially debilitating for different students, but also that students' performances are differentially facilitated by achievement-related anxiety. Studying together with others should, therefore, be more debilitating for some students and more facilitating for others than studying alone.

Variations in the procedure for reporting performance on tasks may also produce varying amounts of arousal. Reporting of group performance

(a cooperative setting) should lead to less anxiety than when individual performance is to be reported (competitive settings). Therefore, the present author predicts that differences in test performance between students who are facilitated by anxiety and those who are debilitated will be greater in the competitive setting than in the cooperative setting.

Method

Design

Six treatment conditions were used in a 2 x 3 between subjects analysis of variance design. The first factor was defined by the levels studying together and studying alone. The second factor, grading procedure, was manipulated by instructing Ss that they would be given credit on the basis of (1) their test score combined with a partner's score (cooperative), (2) their score in comparison with a partner's score (competitive), or (3) their score in comparison with the scores of all others taking the test (normative). The dependent variable was the S's score on a fifteen minute, short-answer test over material presented in a lecture during the first twenty minutes of the experimental session.

In order to test the hypotheses of differential effects of treatments on learners with different dogmatism, facilitating anxiety, and debilitating anxiety scores, comparisons of regression weights were made across treatment conditions.

Subjects

One hundred seventy-four students from an introductory, undergraduate educational psychology course at The Pennsylvania State University

volunteered for participation in the experiment. The Ss were given extra grade credits in the course for participating in the experiment. All except ten of the Ss had taken the Dogmatism Scale, Form E (Rokeach, 1960) and the Achievement Anxiety Test (Alpert & Haber, 1960) in a battery of tests and questionnaires administered six to eight weeks prior to their participation in the experiment. Means and standard deviations for the Dogmatism and Anxiety Scales are presented for the treatment conditions in Appendix 1.

The ten Ss who had not taken the test battery were eliminated from analysis of the data. All ten made substantially higher scores on the criterion test than the other Ss and five of the ten were included in one of the six treatment conditions. The E believes that these ten Ss were, as a group, more highly motivated than the other Ss, because they had not received the credit for participation in the original test battery and the present experiment provided more course credit for those who performed well in the experiment.

Materials

An eighteen minute, taped lecture on the lives of William James, G. Stanley Hall, and John Watson served as the stimulus materials. The biographical information included important dates, names, and places, and excluded information about their contributions to psychology, topics with which many Ss would have had some familiarity. With some minor editing, the text for the lecture was taken from Watson (1963, pp. 320-330 and 385-390) on James and Watson and from Boring (1957, pp. 518-521) on Hall.

The criterion measure was a test of 60 items, 20 concerning each of the three psychologists. The items required recall of dates, names

of individuals or books, and places. The 60 items used were chosen from a pool of 72 on the basis of pilot work that indicated some of the original items were correctly answered by almost all or almost none of the Ss.

Procedure

The basic procedure involved presentation of the eighteen minute lecture, a ten minute study period, and the test which had a fifteen minute time limit.

The Ss were first instructed that the experiment's purpose was to determine how well students learned from tape recorded lectures, and that the lecture would concern the details of the lives of three noted educational psychologists. They were then told that after the lecture they would be able to use a pencil and paper to note what they remembered in preparation for a test on the details of dates, names, and places given in the lecture.

Grading procedure manipulation. The Ss were then informed that the credit they would receive for participation in the experiment would be determined by the score they made on the test. Ss in the cooperative condition were told that credits received would be determined by the combination of their score with that of another S. In the competitive condition Ss were told that credit would be determined by comparing their score with that of one other S. And Ss in the normative condition were told that credit would be assigned on the basis of a comparison with all Ss' scores. All Ss were told that the credit they received would be posted outside the experiment room within a week.

Study procedure manipulations. The lecture tape was then played. After the lecture pencils and paper were distributed and Ss were told

that they had ten minutes to study for the test. Ss in the study together condition were told that they could exchange information, on 5 x 8 cards provided them, with the person seated in the study area opposite them. (If the Ss were in either the cooperative or competitive grading condition, they were also informed that their information exchange partner was the person with whom they had been paired in determining the number of credits they would receive.) In the study alone condition Ss were given no further instructions.

After the study period the 60 item criterion test was administered and Ss were informed that there was a fifteen minute time limit. When Ss said they were finished before the fifteen minute limit, they were asked to review their answers and unanswered questions until the time had lapsed.

The experimental setting was a twelve by eighteen foot room with a four by five foot table in the middle. The table was partitioned into six study areas by eighteen inch high partitions, which overlapped the ends of the table by six inches. Three chairs were placed on two opposite sides of the table. The Ss sat in the chairs at the table during the entire session, thereby preventing a view of the other Ss. Information was exchanged on the 5 x 8 cards through one inch high slots cut into the bottom of the partition separating study areas directly opposite one another.

The procedure for assigning subjects to treatment conditions was as follows. If an even number of Ss arrived for the session, they were assigned to the study together condition. If an odd number of Ss arrived, they were assigned to the study alone condition. Grading procedure conditions were assigned to sessions in an alternating

fashion, with the constriction that all conditions should maintain an approximately equal number of Ss.

Results

A 2 x 3 factorial analysis of variance for unequal N's using the harmonic means (Winer, 1962, pp. 222-4) was computed on the criterion test scores. There were no main effects, the means for both study procedures and all three grading procedures being approximately equal. The interaction between study procedure and grading procedure was significant, with an F (2,158) = 3.44, $p < .05$. The nature of this interaction, which is contrary to the one expected, is depicted in Figure 1.

A comparison of the regression of Dogmatism Scale scores on the experimental criterion test between the study together and study alone conditions yielded a non-significant F (1,160) = 2.22. The more detailed comparison of the regression of dogmatism on criterion performance in the six treatment groups resulted in an F (5,152) = 0.72. The related correlation coefficients, presented in Table 1, are in the predicted direction for the study procedure comparison, though a different configuration for the six treatment conditions was predicted.

None of the comparisons of the regression of anxiety scores on the criterion scores yielded F's greater than 1.10. The relevant correlations for Debilitating Anxiety and Facilitating Anxiety are presented in Appendix 2.

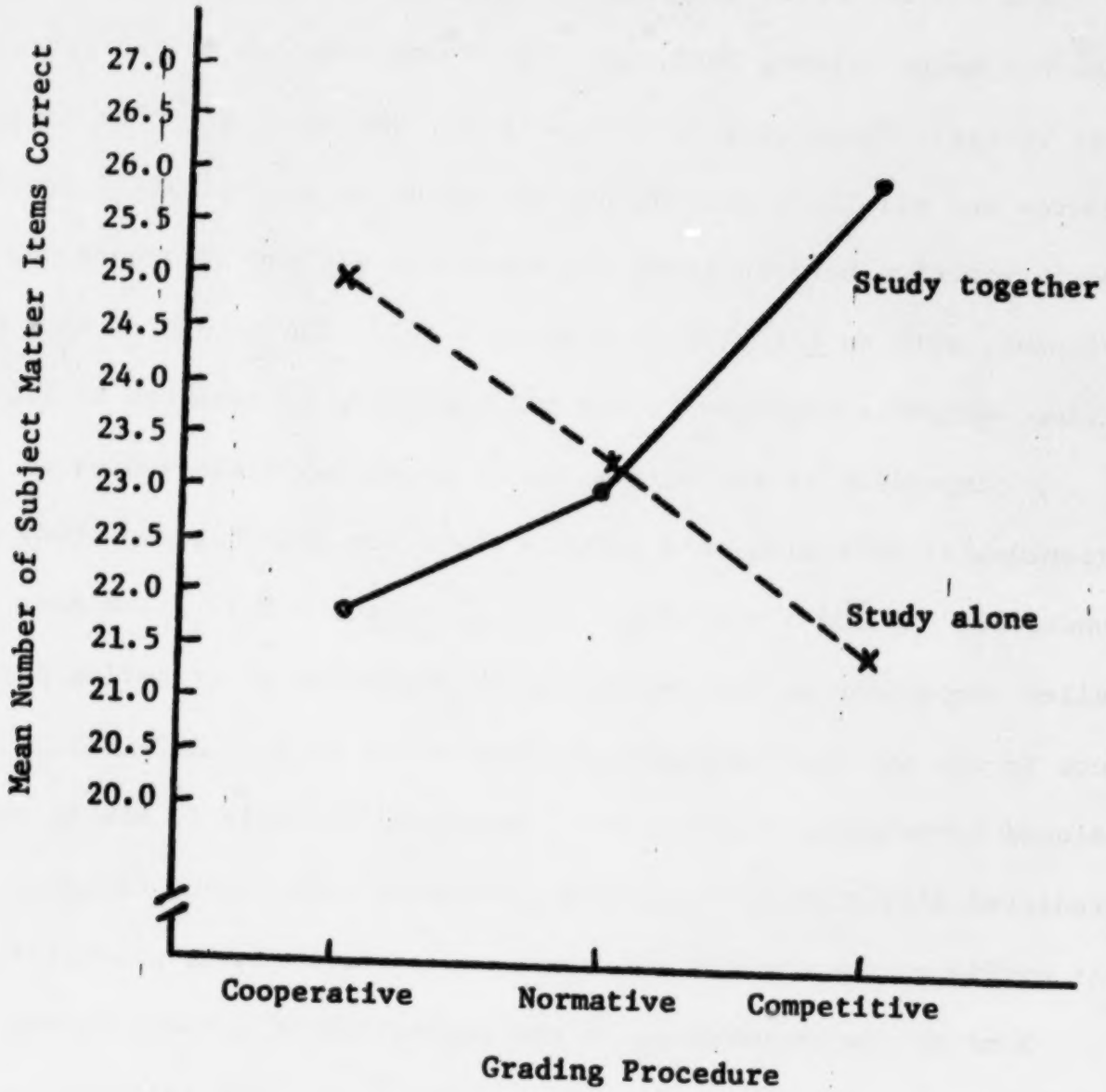


Figure 1. Amount of subject matter recalled as a function of study procedure and grading procedure.

Table 1

Correlation Between Dogmatism and Amount of Subject Matter Recalled in the Six Treatment Conditions and in the Two Study Conditions

Grading Procedure	Study Procedure	
	Study alone	Study together
Cooperative	+ .03	- .22
Normative	+ .08	- .46*
Competitive	+ .09	- .11
Grading procedures combined	+ .01	- .25*

* $p < .05$.

Discussion

The nature of the interaction between study procedures and grading procedures is in marked contrast to the one expected from the rationale based on pooling of information. The original conceptualization implied that pooling, and therefore criterion scores, would increase as the grading procedure lead to more cooperation; no differences were expected among grading procedures in the study alone condition.

To assess the methodological sufficiency of the information exchange procedure for pooling, an analysis of the 5 x 8 information exchange cards of a sample of one-third of the Ss was made. The analysis revealed that information was requested and provided by 82% of the pairs. However, subject pairs in all three grading conditions were approximately equal in information exchange, indicating that the manipulation of the grading procedure did not seem to have an impact on amount of information exchanged. Thus, one may assume that the exchange of information was serving another function, instead of or in addition to pooling. The other confusing aspect of the interaction found is the differences among the grading procedure means in the study alone condition. Differences among those means would seem to indicate differences in the effects of grading procedures in the absence of opportunity for information exchange.

There is at least one interpretation of the present findings that could serve as a basis for additional empirical investigation. This interpretation is based on the assumptions that (1) the grading procedure differences resulted in differences in the Ss' estimations of their chances of success, which, in turn, affected their amount of effort, and (2) the study together procedure had its impact, not so

much by effecting pooling, but in allowing for sounding out of the capabilities of one other S, which, in turn, affected the estimated chances for success.

In the study alone condition, Ss could obtain no definitive estimate of what other Ss remembered from the lecture. In the absence of such information the normative grading condition Ss might have used past experience in normative grading situations to estimate chance of success in the present situation; therefore, his degree of effort in the present task would have been the same as it usually was. The cooperative and competitive grading procedures were relatively novel to the student and might have resulted in a change in his perceived chances of success. In both the cooperative and competitive situations, the person's final grade was dependent on the performance of one other person. The probability of success was not assessable, since the unknown other person might score well or poorly or average. In general, one would expect a lower degree of effort under pure chance circumstances, and Ss in the competitive condition would seem to conform to this expectation. However, Ss in the cooperative condition produced higher scores than those in both other grading conditions; the greater effort in the cooperative condition may have been the result of taking responsibility for the success of the other, perhaps less capable, S.

In the study together procedure, Ss could obtain an estimate of how much one other student remembered. One may assume that on the average the other student remembered some things S did not. Of the three grading conditions, the normative one would seem to be least affected by this information, since he was to be compared with all Ss. The cooperation condition S might have been satisfied with his partner's

knowledge, and therefore would believe that his chances of success were sufficient without even usual effort on his part. However, the reaction of the S in the competitive condition might have been that greater effort was required to maintain at least partly with the other.

While the interpretation given above is post hoc and requires additional empirical support, the implication has not been considered formally in group problem solving research. In any achievement situation the capabilities of the individuals are manifest as a function of the individuals' perceptions of chances of success. When chances of success are dependent on the performances of others - as they are in many life situations, the interchanges between the individuals will affect the perceived chances of success, in addition to allowing for information exchange. The extent to which the sounding out process predominates over the information exchange process should be a topic of further investigation; since a primary purpose of group work is the pooling of skills and knowledge, the sounding out process would seem to be an interfering factor. Research on aspects of the situation that reduce (or adequately provide for satisfaction of) the need for sounding out the other persons would be valuable.

The individual difference variables predicted to be related to the issue of studying together and alone under various grading conditions were not related to criterion performance. The nonsignificance of the predicted relationships may have been the result of the different set of factors outlined above. Dogmatism was predicted to be negatively related to performance in the study together procedure because high dogmatics were assumed to profit less from information provided by their peers. If the sounding out process predominated over the pooling

processes, dogmatism would not have been a salient learner difference related to performance. However, it should be noted that the one correlation between dogmatism and test performance that was significant was in the study together, normative grading treatment, the treatment interpreted as least affected by the sounding out process.

The unsupported predictions concerning differential effects of anxiety on test performance among the various treatment conditions may have been the result of inadequate procedures and/or inappropriate measures. The prediction concerning differences between study alone and study together conditions was based on Zajonc's (1965) hypothesis that work in a social setting is more activating than working alone. In the present research, all Ss shared an experimental setting with other Ss and the credits all Ss obtained were posted for all to view; the additional social feature of a ten minute information exchange with another student may not have made the study together setting sufficiently more social than the study alone condition. Also, the use of Alpert and Haber's (1960) scales for the measurement of anxiety arising from different grading procedures may be inappropriate. The validity of the scales may be limited to the usual situations of normative evaluation, since that was the only condition in the present study to have correlations that were significantly different from zero.

In conclusion, the results of the present study were unexpected. The post hoc interpretation of the interaction between study conditions and grading procedures requires additional empirical support. The additional research should include interim measures of the Ss' perceived chances of success and should manipulate variables predicted to modify the sounding out process postulated for study together conditions. The

important individual difference variables in the further research would be differences in reaction to different levels of perceived chances of success, such as these developed by Atkinson (1966).

References

Alpert, R., & Haber, R. N. Anxiety in academic achievement situations. Journal of Abnormal and Social Psychology, 1960, 61, 207-215.

Atkinson, J. W. A theory of achievement motivation. N.Y.: Wiley, 1966.

Boring, E. G. A history of experimental psychology. (2nd ed.) New York: Appleton Century, 1957.

Deutsch, M. An experimental study of the effects of cooperation and competition on group processes. Human Relations. 1949, 2, 199-232.

Ehrlich, H. J., & Lee, D. Dogmatism, learning, and resistance to change. Psychological Bulletin, 1969, 71, 249-260.

Powell, F. A. Open- and closed-mindedness and the ability to differentiate source and message. Journal of Abnormal and Social Psychology, 1962, 65, 61-64.

Rokeach, M. The open and closed mind. New York: Basic Books, 1960.

Watson, R. D. The great psychologists: From Aristotle to Freud. New York: Lippincott, 1963.

Winer, B. J. Statistical procedures in experimental design. New York: McGraw-Hill, 1962.

Zajonc, R. B. Social facilitation. Science, 1965, 149, 269-274.

Zander, A., & Wolfe, D. Administrative rewards and coordination among committee members. Administrative Science Quarterly, 1964, 9, 50-69.

Appendix 1

Means and Standard Deviations for the Two Achievement Anxiety Subscales and the Dogmatism Scale for Each Treatment and Treatments Combined

Treatment		Debilitating Anxiety		Facilitating Anxiety		Dogmatism	
	(N)	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD
Cooperative, study together	(27)	28.0	4.53	24.4	4.04	139.4	21.06
Competitive, study together	(27)	27.7	4.90	25.0	3.85	139.9	16.76
Normative, study together	(26)	28.0	5.93	24.7	4.66	141.1	19.33
Cooperative, study alone	(29)	25.9	6.08	25.3	4.94	135.0	14.91
Competitive, study alone	(28)	29.0	6.03	21.8	5.24	135.8	16.28
Normative, study alone	(27)	29.9	6.30	24.0	3.20	144.1	16.88
Cooperative	(56)	26.9	5.45	24.9	4.51	137.1	18.11
Competitive	(55)	28.4	5.50	23.4	4.84	137.8	16.49
Normative	(53)	29.4	6.08	24.3	3.97	142.6	18.01
Study together	(80)	28.2	5.10	24.7	4.15	140.1	18.90
Study alone	(84)	28.2	6.30	23.7	4.75	138.2	16.35
Treatments combined	(164)	28.2	5.73	24.2	4.48	139.1	17.61

Appendix 2

Correlations Between the Two Achievement Anxiety Subscales and Amount of Subject Matter Recalled in the Six Treatment Conditions and in the Two Study Conditions

Debilitating Anxiety

Grading Procedure	Study Procedure	
	Study alone	Study together
Cooperative	- .12	- .10
Normative	- .32*	- .47*
Competitive	- .19	+ .12
Grading procedures combined	- .23	- .16

Facilitating Anxiety

Grading Procedure	Study Procedure	
	Study alone	Study together
Cooperative	+ .11	+ .20
Normative	+ .27	+ .61**
Competitive	+ .11	- .05
Grading procedures combined	+ .09	+ .26*

* $p < .05$
** $p < .01$

Summary

Reliabilities of Six Personality Measures Used in the Instructional Strategies Research Project

Nicholas M. Sanders and Paul D. Weener

Technical Problem

The reliability of a measure sets a limit on the degree of relationship that measure may exhibit with any other variable. Many studies in the "Instructional Strategies" research project included investigations of the relationship of one or more of six personality variables to instructional treatment outcomes. Therefore, an adequate interpretation of the findings of those studies should include a consideration of the reliabilities of the measures used. The purpose of the present report is to present reliability data as an aid in interpretation of the findings of the individual studies utilizing the six personality measures.

General Methodology

Rotter's (1966) Internal-External Scale, Rokeach's (1960) Dogmatism Scale, Form E, Crowne and Marlowe's (1964) Social Desirability Scale, Budner's (1962) Intolerance of Ambiguity Scale and the Facilitating and Debilitating Anxiety Subscales of Alpert and Haber's (1960) Achievement Anxiety Test were administered in large group testing sessions during three school terms. Internal consistency estimates of reliability were computed each term for each scale. During the Winter term, 1970, a group of students repeated the battery, thus allowing for computation of a stability estimate of reliability for each scale.

Technical Results

Of the six scales the Debilitating Anxiety Scale alone was evaluated as having completely adequate reliability, while the Facilitating Anxiety Scale was judged to have a reliability low enough to warrant special caution in interpretation of results of individual project studies. The Intolerance of Ambiguity Scale, the Social Desirability Scale and the Internal-External Scale had adequate or high stability, but the internal consistency coefficients indicated that the scales were not homogeneous. In contrast, the Dogmatism Scale was homogeneous, though test-retest results suggest the scale may be relatively unstable.

Reliabilities of Six Personality Measures Used in the Instructional Strategies Research Project

Nicholas M. Sanders and Paul D. Weener[1]

Many of the individual studies in the Instructional Strategies research project involved an analysis of the relationship of one or more of six personality measures to criteria relating to experimental instructional treatments. The personality measures, the Debilitating Anxiety and Facilitating Anxiety Scales (Alpert & Haber, 1960), the Intolerance of Ambiguity Scale (Budner, 1962), the Social Desirability Scale (Crowne & Marlowe, 1964), the Dogmatism Scale (Rokeach, 1960), and the Internal-External Scale (Rotter, 1966), were used in order to explore the presence of hypothesized differences in the effects of the treatments as a function of learner differences in these personality characteristics. The reliabilities of these measures in the population from which subjects were drawn for particular studies is of concern since most analyses involved the use of individual subjects' scores, instead of groupings of subjects having similar scores (as in comparisons of subjects attaining high scores with those attaining low scores).

The present report serves three functions. First, reliability coefficients for the population from which samples were drawn for individual studies aid in the interpretation of individual difference and

1 - The authors acknowledge the contributions of Charles Schultz, who coordinated data collection, and Ovid Tzeng, who aided in analysis of the data.

criterion relationships revealed in those studies, because the degree of relationship between a measure and another set of scores is limited by the reliabilities of each set. Second, internal consistency reliability data are presented for each of three quarters of the school year, 1969-1970. A comparison of the three coefficients for the different terms should provide information on possible differences among subjects at different times during the year and on possible differences resulting from a change in administration procedures during the third term. Third, the report presents reliability coefficients obtained in original standardization work on the measures, allowing a comparison of the measures' reliabilities for the present population with those for the original standardizing population.

Method

Design

Data were analyzed separately for each of three school terms to provide information about any variation in the characteristics of the tests as a function of term. Internal consistency estimates of reliability were obtained for each of the terms, and a stability estimate was provided by retesting a smaller group of Ss during one of the terms.

Subjects

A total of 1,899 students volunteered to participate in the study over three quarters of the school year, 1969-1970. The Ss were students in an introductory educational psychology course at The Pennsylvania State University, and they received extra grade credit in the course for their participation in the study.

Measures

The personality measures used were the I-E Scale, measuring internal or external control of reinforcements (Rotter, 1966), the Scale of Tolerance-Intolerance of Ambiguity (Budner, 1962), the Social Desirability Scale (Crowne & Marlowe, 1964), the Achievement Anxiety Test, yielding a facilitating and a debilitating anxiety score (Alpert & Haber, 1960), and the Dogmatism Scale, Form E (Rokeach, 1960).

Procedure

Fall term. The battery of five questionnaires was administered in three large group sessions (80 to 200 Ss participating in each), two smaller group settings of up to 20 Ss who could not attend the large group sessions, and about five individual sessions. Testing was completed by the end of the third week of the ten week term.

A general introduction for each session included an overview of the purpose of the test battery and an assurance of anonymity of individual performances. The E then distributed a booklet with the questionnaires presented in the following order: (1) I-E Scale, (2) Intolerance of Ambiguity Scale, (3) Social Desirability Scale, (4) Achievement Anxiety Test, and (5) the Dogmatism Scale. Ss responded on a ten-alternative, multiple-choice answer sheet. No time limit was set, and Ss were told they could leave after they had completed all the questionnaires.

Winter term. The procedure used in the Fall term was repeated. However, in addition, two retest sessions were held approximately three weeks after the original testing to enable computation of a test-retest reliability. In one session 72 Ss repeated the I-E Scale, the Social

Desirability Scale, and the Achievement Anxiety Test. In the other session 59 Ss were retested on the Intolerance of Ambiguity and Dogmatism Scales.

Spring term. Several procedural modifications were made in the Spring term battery administration. First, there were no large group sessions; instead Ss took the battery in groups of an average size of 30 and with a maximum of 40. The reduction in size enabled a more adequate monitoring by E. Second, the Remote Associates Test, Form 1 (Mednick & Mednick, 1967) was added, and the I-E Scale and the Social Desirability Scale were not included. The new administration sequence of the scales was as follows: (1) Remote Associates Test, (2) Dogmatism Scale, (3) Intolerance of Ambiguity Scale, and (4) Achievement Anxiety Test. Third, all instructions were printed on the booklet and were read to Ss. Finally, all tests were given a time limit: 25 minutes for the Remote Associates Test, 20 minutes for both the Dogmatism Scale and the Intolerance of Ambiguity Scale together, and ten minutes for the Achievement Anxiety Test.

Results

Means, standard deviations, and reliability coefficients are presented in Tables 1 through 5. Notes are made for each table to indicate the method of reliability coefficient computation. Each scale, with the exceptions of the Achievement Anxiety Test subscales, is presented in a separate table to allow for comparison of means, standard deviations, and reliabilities among the three school terms and between the present study and the original standardization studies on the scales. Data for the Achievement Anxiety Test subscales of Debilitating and Facilitating Anxiety are presented together in Table 1.

Table 1

Reliabilities, Means, and Standard Deviations

for Debilitating Anxiety and Facilitating Anxiety

Subscales of the Achievement Anxiety Test

Type of Reliability and Study	N	Debilitating Anxiety $\bar{X}$	SD	$r_{xx'}$	Facilitating Anxiety $\bar{X}$	SD	$r_{xx'}$
Internal Consistency							
Fall, 1969	588	26.9	6.29	.83[a]	24.6	4.41	.65[a]
Winter, 1970	617	26.8	5.83	.79	25.0	4.74	.63
Spring, 1970	685	26.4	5.86	.82	24.8	4.41	.67
Stability (Test-retest)							
Winter, 1970							
Original		27.0	6.84		25.4	4.86	
	72			.82			.65
Three week interval		27.0	7.41		26.7	4.60	
Albert & Haber (1960)[b]							
Original	379	26.3	5.33		27.3	4.27	
Ten week interval	40	c	—	.87	—	—	.83
Eight month interval	40	—	—	.76	—	—	.75

a - The Hoyt analysis of variance method was used to estimate internal consistency.
b - The Alpert and Haber (1960) study included only males.
c - Not given in study.

Discussion

In the following paragraphs each of the six personality measures will be discussed separately. The discussion includes a general evaluation of the adequacy of the measure's reliability for use in the individual studies included in the "Instructional Strategies" projects, a comparison of the internal consistency coefficients of the measure over the three school terms, and a comparison of the reliability estimates obtained with those reported from the original standardization research.

The Debilitating Anxiety subscale of the Achievement Anxiety Test had both high internal consistency and high stability. The reliability of the scale, therefore, should have been sufficient to allow for high correlations with any other variables predicted to be related. Also, fluctuations of the internal consistency were small among the three school terms. Though no internal consistency coefficients were presented for the subscale by Alpert and Haber (1960), the test-retest reliability found in the present study was only slightly lower than those reported by Alpert and Haber.

In contrast to the Debilitating Anxiety scale, the Facilitating Anxiety scale seemed to be less homogeneous and less stable. The obtained coefficients, while not indicating a totally inadequate scale, do indicate that findings of a lack of relationship between facilitating anxiety and experimental criteria may have been the result of the low reliability of the Facilitating Anxiety scale for the population from which subjects were drawn. As with the Debilitating Anxiety scale, there were only small fluctuations in the internal consistency across the three school terms. In comparing Alpert and Haber's (1960) test-retest correlations with the one obtained in the present study, one

Table 2

Reliabilities, Means, and Standard Deviations for

Intolerance of Ambiguity Scale

Type of Reliability and Study	N	$\bar{X}$	SD	$r_{xx'}$
Internal Consistency				
Fall, 1969	596	44.7	7.09	.50[a]
Winter, 1970	618	44.7	7.56	.54
Spring, 1970	685	42.5	8.13	.63
Budner (1962)[b]	30	30.9	10.13	.62
Stability				
Winter, 1970				
Original	59	44.8	8.40	.73
Three week interval		45.0	8.58	

a- The alpha coefficient was used to compute internal consistency

b- Budner used a seven point scale for each item, while the present research utilized a six point scale. Greater comparability of descriptive statistics may be attained by assuming that the item responses Budner obtained may be transformed into the six point score by multiplying each item response by a constant of 6/7; the resulting mean and standard deviation would be 43.6 and 8.68 respectively, and the reliability would be relatively unaffected since item variances would be transformed also.

Table 3

Reliabilities, Means, and Standard Deviations

for the Social Desirability Scale

Type of Reliability and Study	N	$\overline{X}$	SD	$r_{xx'}$
Internal Consistency				
Fall, 1969	596	13.9	5.38	.75[a]
Winter, 1970	618	13.0	5.55	.57[a]
Crowne & Marlowe (1964)	76	___	___	.88[b]
Stability				
Winter, 1970				
Original	72	13.7	5.39	.86
Three week interval		13.5	5.74	
Crowne & Marlowe (1964)	57	___	___	.88
Additional Norms				
Crowne & Marlowe (1964)				
Ohio State Males	666	15.1	5.58	
Ohio State Females	752	16.8	5.50	
Northwestern Males	100	11.6	5.26	
Northwestern Females	86	13.5	4.75	
U. Washington Males	110	14.4	5.62	

a - An alpha coefficient was used to compute the internal consistency.
b - Internal consistency was computed by Kuder-Richardson, Formula 20 method.

finds a considerable difference, with the Alpert and Haber report indicating a much greater stability than the present study. Since Alpert and Haber's subjects were males and the present study included both males and females, one hypothesis for the differences in reliabilities is that the scale is less reliable for the female population.

The Intolerance of Ambiguity Scale is the least internally consistent test used in the battery. However, Budner (1962) pointed out that the construct being measured by the scale is posited to be a complex, multi-dimensional one; researchers using the scale in their studies should base their predictions on a conceptualization that acknowledges the multi-dimensionality of the scale. Since the scale's reliability is not appropriately assessed by the internal consistency, the stability of the scale becomes more important as the reliability estimate. The stability of the scale is only moderate (.73) in the present study. A comparison of internal consistency coefficients for the three school terms reveals some fluctuation, with the Spring term (during which administration procedures were more rigorous and standardized) being the highest. Also, the Spring term internal consistency is more comparable than the other terms to Budner's own results.

Crowne and Marlowe's (1964) Social Desirability Scale was highly stable, though internal consistency estimates of reliability varied between the Fall and Winter terms, making any overall evaluation of reliability difficult. Since the administration procedures during Fall and Winter were the same, only a difference in the population could have caused the difference in internal consistency. The author has no hypothesis to offer for the difference. The comparison of the present results with Crowne and Marlowe's (1964) reveals comparable stability

Table 4

Reliabilities, Means and Standard Deviations

for the Dogmatism Scale, Form E

Type of Reliability and Study	N	$\overline{X}$	SD	$r_{xx'}$
Internal Consistency				
Fall, 1969	596	135.5	18.39	.79[a]
Winter, 1970	614	134.7	22.56	.86[a]
Spring, 1970	685	128.6	19.38	.80[a]
Rokeach (1960)[c]				
Ohio State I	22	142.6	27.6	.85[b]
Ohio State II	28	143.8	22.1	.74[b]
Ohio State III	21	142.6	23.3	.74[b]
Ohio State IV	29	141.5	27.8	.68[b]
Michigan State	89	———	———	.78[b]
Stability				
Winter, 1970				
Original	59	140.9	18.48	.64
Three week interval		134.4	26.36	
Rokeach (1960)[c]				
Original	58	141.3	28.2	.71
Five to six months interval		143.2	27.9	

a - Internal consistency was computed using the alpha coefficient.
b - Internal consistency was computed using an odd-even item split correlation with a correction for total test reliability by the Spearman-Brown procedure.
c - Rokeach's response form for items on the Dogmatism Scale was a seven-point scale; a six point scale was used in the present study. Greater comparability of descriptive statistics may be attained by assuming that item responses Rokeach obtained may be transformed into the six point scale by multiplying each item response by a constant of 6/7; thus, a mean of 142 and a standard deviation of 25 on the seven point scale would be 121.7 and 21.11, respectively, on the six point scale. The reliability coefficients would be unaffected by the transformation.

Table 5

Reliabilities, Means, and Standard Deviations

for Internal-External Scale

Type of Reliability and Study	N	$\bar{X}$	SD	$r_{xx'}$
Internal Consistency				
Fall, 1969	596	13.5	4.07	.71[a]
Winter, 1970	618	13.1	3.98	.69[a]
Rotter (1966)				
I	100	___	___	.73[b]
II	400	___	___	.70[b]
Stability				
Winter, 1970				
Original	72	10.2	3.88	.91
Three week interval		10.0	4.03	
Rotter (1966)				
One month interval	60	___	___	.72
Two month interval	117	___	___	.55
Additional Norms				
Rotter (1966)				
Ohio State	1180	8.3	3.97	
Kansas State	113	7.7	3.82	
Univ. of Conn.	303	9.2	3.88	

a - Internal consistency was computed using the alpha coefficient.
b - Internal consistency was computed using the Kuder-Richardson procedure.

coefficients, even though those authors' internal consistency coefficient was much greater than those in the present study.

The homogeneity of the Dogmatism Scale was high in the present study, but the test-retest correlation indicates that the scale may be unstable for the population from which subjects were drawn for the individual studies. The relative instability of the scale is critical to those studies since the studies using the dogmatism variable were carried out at least three weeks after the Ss completed the Dogmatism Scale. Internal consistency coefficients fluctuated relatively little, especially when compared with the variations presented by Rokeach (1960). The comparison of Rokeach's five to six month stability with the present study's three week stability reveals a great difference. However, an examination of means and standard deviations of the sample on which the present stability study was based suggests that (1) the Ss may not have been representative of the Winter term population and (2) the Ss may have been especially unstable in their dogmatism.

The Internal-External Scale was very highly stable, though of only moderate internal consistency. Aside from a precaution that the scale may measure a slightly heterogeneous construct, there is no reason to believe results of studies would be affected by lack of reliability of the scale. Internal consistency was comparable across school terms and was similar to that presented by Rotter (1966). The stability found in the present study was much higher than that reported by Rotter; though population differences between colleges may account for the difference in stability, Rotter also notes that the original standardization testings were in group settings for the first testing and in individual settings for the retest.

References

Alpert, R., & Haber, R.N. Anxiety in academic achievement situations. Journal of Abnormal and Social Psychology, 1960, 61, 207-215.

Budner, S. Intolerance of ambiguity as a personality variable. Journal of Personality, 1962, 31, 29-50.

Crowne, D.P., & Marlowe, D. The approval motive: Studies in evaluative dependency. New York: Wiley, 1964.

Mednick, S. A., & Mednick, M. T. Remote Associates Test. Boston: Houghton Mifflin, 1967.

Rokeach, M. The open and closed mind. New York: Basic Books, 1960.

Rotter, J. B. Generalized expectancies for internal versus external control of reinforcement. Psychological Monographs, 1966, 80, No. 1 (Whole No. 609), 1-28.

Summary

The Effects of Uncertainty, Confidence, and Individual Differences on the Initiation and Direction of Information-seeking Behaviors*

Study Director: Charles B. Schultz

Advisor: Francis J. Di Vesta

Technical Problem

The present study was an investigation of the conditions that induce individuals to seek and acquire information (epistemic curiosity). The initiation of epistemic curiosity has been attributed, in large part, to the amount of uncertainty produced by stimuli which elicit competing response alternatives (Berlyne, 1962). Uncertainty and the consequent epistemic curiosity were assumed to be heightened when the number of competing responses is increased or when the responses are of equal or close-to-equal strength. Moreover, the drive-like state of curiosity is reduced by the acquisition of knowledge which reduces response competition. Thus, the model of curiosity used in the present investigation is one of drive reduction.

This study also explored the conditions that induce individuals to seek and acquire discrepant information, that is, information

---

* An earlier progress report entitled "Uncertainty: A Basis for Instructional Strategies That Initiate and Direct Information-Seeking Behaviors" was included in the January, 1970, Semi-Annual Report. The present summary is of a Ph.D. dissertation conducted under this contract. The dissertation will also appear as a separate technical report.

inconsistent with beliefs they already hold. According to Festinger's (1957) early notions of cognitive dissonance, knowledge that information is inconsistent with existing beliefs comprises a set of conflicting cognitions. Since the resulting dissonant state is psychologically disturbing for the organism, discrepant information is avoided. In order to explain instances in which individuals have sought or at least failed to avoid discrepant information, Festinger (1964) modified his earlier formulations by suggesting that individuals may be receptive to discrepant information when it is useful and when they are sufficiently confident of their ability to refute the counter-arguments posed by the discrepant information.

The present investigation assumed that discrepant information may be sought for its intrinsic utility. This condition occurs when an existing belief is suddenly found to compete with another that also appears valid. As a result, the two alternatives may have close-to-equal strengths. Under these circumstances an individual may actually seek information regarding the discrepant alternative to reduce the equality of the competing responses.

In typical selective exposure experiments, Ss are not in the position of avoiding dissonance, but of reducing it. Under these circumstances, the Ss may select and examine discrepant information if they are confident they can counter arguments posed by the discrepant material. In addition, since confidence typically has been induced by informing Ss that their responses are correct (high confidence) or are incorrect (low confidence) according to standards set by an expert, an authority, or by E, it was reasoned that the effectiveness of the confidence manipulation would be modified by the individual predisposition

to accept feedback attributed to an authority. Accordingly, dogmatic persons who experience low confidence must seek discrepant information to be consistent with authority beliefs, while dogmatic persons who experience high confidence must seek congruent information to be consistent with authority beliefs. Other individual differences may affect the direction of information-seeking. Discrepant information may hold less utility for individuals who fail to generate response competition or who avoid ambiguous situations than for those who are "quick" to generate response competition or who are attracted to ambiguous situations. Therefore, an individual's tendency to be subjectively certain or intolerant of ambiguity affects the amount of discrepant information he seeks.

Based on this rationale, it was hypothesized that uncertainty is directly related to the examination and acquisition of knowledge about the general experimental topic and of the position which is discrepant with the one the individual holds. It was also expected that experimentally induced confidence and personality traits such as subjective certainty, intolerance of ambituity, and dogmatism are inversely related to the seeking and acquisition of discrepant information. A final hypothesis was that confidence is inversely related to the seeking and acquisition of discrepant information for closed-minded persons and unrelated for open-minded persons.

General Methodology

Two experiments were conducted to test these hypotheses. The first investigated the effects of uncertainty (Experiment I) on information-seeking; the second examined the effects of confidence and

personality differences on information-seeking at one level of uncertainty (Experiment II).

Tests designed to measure dogmatism, intolerance of ambiguity, and an Uncertainty Scale specifically developed for this experiment were administered to students in education courses, all of whom were potential Ss for the study, several weeks before the experiments were conducted.

In both experiments Ss were told that the experimenters were preparing instructional materials on the topic of attitude change. It was explained that the S's task was to examine a pair of slides containing information on attitude change and to select what he considered the more interesting member of the pair. The two slides in the pair were projected simultaneously by two carousel projectors. The experimental materials included 14 slide pairs in which the two members were identical descriptions of the general procedures and results of an experiment on attitude change. In the 16 remaining pairs, a slide containing information which was congruent with S's beliefs was paired with a slide containing information which was discrepant with his beliefs. These critical pairs of slides were balanced in form, length, and content.

Dependent measures of information-seeking and acquisition included the time spent examining slides chosen as more interesting, ratings of interest, and scores on a multiple-choice test administered at the completion of the experiment. Measures of selective exposure included the number of discrepant slides chosen, the time spent examining discrepant slides, sub-test scores for retention of information about the

congruent and discrepant positions, and self-reports of interest in the two positions.

The treatments in Experiment I consisted of the manipulation of three levels of uncertainty: Incongruity, Doubt, and Certainty. In the Incongruity Condition, Ss were shown evidence supporting dissonance theory that contradicted their position. The Doubt Condition consisted of presenting Ss with both supporting and contradictory evidence. In the Certainty Condition, Ss were only shown evidence supporting reinforcement theory that agreed with their position. Finally, no evidence was presented to Ss in the Control Condition. These treatments imply a completely randomized design with three experimental groups (Incongruity, Doubt, and Certainty) and one control.

All Ss in Experiment II received the Doubt instructions administered in Experiment I. In addition, Ss in two of the groups were administered a test that was intended to measure "intuitive understanding of attitude change." The Ss in one of these groups were told that their responses placed them in the 93rd percentile, thereby inducing the condition of High Confidence. The Ss in the other group were told that their scores placed them in the 11th percentile, thereby inducing the condition of Low Confidence. The Doubt Condition employed in Experiment I served as the control for Experiment II.

These treatments imply a completely randomized design with two experimental groups (High and Low Confidence) and a control. In order to assess the relationship between confidence and dogmatism, a regression analysis was made in which dogmatism was considered the independent variable and measures of discrepant information the dependent variables

for each of the treatment groups. Finally, dogmatism, intolerance of ambiguity, and subjective uncertainty were correlated with measures of selective exposure for Ss in the Doubt Condition.

Technical Results

The Ss who had been exposed to evidence which contradicted their beliefs (i.e., the Incongruity and Doubt manipulations) examined and acquired more information on the experimental topic than Ss who had been exposed to evidence which agreed with their existing beliefs. Presumably, the effect of the discrepant evidence was to strengthen new or subordinate beliefs, thereby sharpening response competition with the consequent arousal of epistemic curiosity. As a result, Ss engaged in epistemic behaviors (observation) which resulted in the acquisition of new information.

The analyses of data obtained on measures of interest and the examination and acquisition of congruent and discrepant information yielded an interaction between levels of uncertainty and type of information sought. Certainty Ss preferred, sought, and acquired congruent information while Incongruity Ss preferred, sought, and acquired discrepant information. The information-seeking behavior of Ss in the Certainty Condition was consistent with expectations based on dissonance theory; namely, information which could increase dissonance was avoided. However, it is difficult to account for the seeking of information which could increase dissonance by Ss in the Incongruity Condition within the context of dissonance theory. The dissonance-increasing behaviors can be explained by the hypothesis of intrinsic utility. According to this notion, the examination and acquisition of information related to the

4

new or subordinate belief suggests that information about that belief was useful for the reduction of response competition and was therefore the focus of epistemic behaviors.

Confidence was unrelated to any of the measures of selective exposures including the DSC and the D/E Ratio, although the confidence manipulation was successfully induced. Analyses of the relationship between dogmatism and confidence revealed a tendency for dogmatic persons to seek and acquire more discrepant information under conditions of Low Confidence (when authorities endorsed discrepant beliefs) than under High Confidence (when authorities endorsed their existing beliefs). These tendencies, however, were not reliable. One reason for the lack of reliability may have been the relatively weak "authority image" projected by E. As a consequence, dogmatic Ss may not have been as influenced by the authority's alleged belief as they otherwise would have been and therefore they did not seek information about the beliefs advocated by the authority.

Dogmatism did not correlate with any of the measures of selective exposure. The failure to obtain the hypothesized inverse relationship between dogmatism and preference for discrepant information may have been due, in part, to the global nature of the dogmatism construct. That is, dogmatism may include components which do not directly relate to the requirements of the experimental task (e.g., authority-orientation, compartmentalization and dichotomization of beliefs). Therefore, the correlation between it and selective exposure was low. In this regard, two task-specific personality differences were found to be reliably related to the acquisition of discrepant information in such a way that

the predispositions to be uncertain and tolerant of amibguity, facilitated learning and the tendency to be certain or intolerant of ambiguity inhibited learning.

Educational Implications

In its present state, much of instructional practice relies on procedures which are based on certainty rather than uncertainty. These include lectures and texts which tend to be highly organized and complete as well as drill-type procedures in which the learner's dominant response is elicited and reinforced. One implication of the present findings is that student learning will be facilitated when it follows the generation of uncertainty. Accordingly, in constructing instructional materials and strategies, the use of open-ended questions, content containing conflicting interpretations, and phenomena which violate the learner's expectations are recommended. These techniques have been included in curriculum projects designed to stimulate student discovery, inquiry, or reflective thought which have utilized uncertainty as a criterion for the selection of instructional topics and as a motivational device sequenced throughout instruction to maintain the learner's explorations.

A second implication of the findings is that uncertainty can be employed to direct the learner's search for new information away from his existing beliefs and thus broaden the scope of his learning. In this regard, incongruity appears to be appropriate as a strategy for implementing instructional objectives which require the learner to focus on information associated with beliefs which contradict those he currently holds. The use of doubt is suggested by the findings to implement objectives which require the learner to "openly" explore

conflicting alternatives or to synthesize information gleaned from various alternatives to form a new generalization.

Finally, Ss who were tolerant of ambiguity and subjectively uncertain acquired more discrepant information than those who were intolerant of ambiguity and subjectively certain. This finding suggests remediational procedures for those who tend to be certain of their responses in problematic situations. These procedures include directing the learners to generate alternate hypotheses and reinforcing the reasonableness of the various alternatives rather than the correctness of a single answer.

Implications for Future Research

The results of the present experiment suggest several lines of investigation. The first concerns the further examination of the theoretical constructs upon which this study was based. One such construct is epistemic curiosity. Although the drive reduction model of curiosity implies that epistemic curiosity is directly related to the acquisition of knowledge, little research has been conducted which attempted to establish this relationship. The present findings are consistent with the drive-reduction interpretation; however, they are not the result of precise manipulations of the determinants of epistemic curiosity. Research of this type would require relatively exact control of the number and relative strengths of the competing responses.

Research of a more applied nature is also suggested by the present findings. Experimental issues of this type include the optimal pacing of uncertainty to maintain the learner's exploration, the effect of teacher questions on maintaining the search and acquisition of knowledge,

and the development remedial programs to generate states of "subjective" uncertainty in learners who are overly certain in problematic situations.

References

Berlyne, D. E. Uncertainty and epistemic curiosity. British Journal of Psychology, 1962, 53, 27-34.

Festinger, L. A Theory of Cognitive Dissonance. Stanford: Stanford University Press, 1957.

Festinger, L. & Carlsmith, J. M. Cognitive consequences of forced compliance. Journal of Abnormal and Social Psychology, 1959, 58, 203-210.

Summary

Recitation Strategies: The Effect of Rates and Schedules of Verbal Responses on Retention

Charles B. Schultz

Technical Problem

The commonly used instructional practice of recitation was assumed to be a stressful condition which some learners have associated with task-relevant responses (facilitators) and others have associated with task-irrelevant responses (debilitators). Thus, recitation strategies may produce rehearsal and coding responses necessary for the transfer of information from short-term to long-term storage for facilitators and produce competing responses which interfere with rehearsal and coding for debilitators. Based on this rationale, an interaction between personality and experimentally induced stress was hypothesized in which retention of facilitators was expected to improve with increased stress and retention of debilitators was expected to decrease with increased stress.

General Methodology

A laboratory experiment was conducted to test the above hypothesis. Groups of six S's viewed slides which contained brief paragraphs describing an experiment on attitude change and then referred to printed versions of the slides to answer recitation questions. Stress was manipulated by varying the rate of verbal responses (7, 3, and 0 responses) and by informing (determined schedule) or not informing (undetermined

schedule) Ss of the questions they would be required to answer during the experimental "lesson." Retention was assessed by a multiple- choice test of information contained on the slides.

Technical Results

The main effects of response rate, response schedule, and personality were all found to be significant. The three and no response rate conditions tended to produce more retention than the seven response rate condition. Facilitators retained more of the information contained on the slides than debilitators. The retention scores of debilitators tended to decrease as response rate increased. However, an interaction between personality and response rate or response schedule was not obtained.

Educational Implications

The findings suggest that moderate rates of recitation and undetermined schedules result in effective retention. In addition, stress-producing strategies such as recitation appear to depress retention of debilitators in relation to facilitators. The relatively poor performance of debilitators may be due to the task-irrelevant responses made by them in stressful conditions. The low retention scores of debilitators, in particular, suggest that the instructor should minimize the stress of recitation strategies and avoid calling on debilitators who appear to retain more by listening than by participating.

Implication for Future Research

In the present study, salient features of the recitation situation were examined for their effect on retention. The differential retention

scores for facilitators and debilitators implies the need for instruction which provides remedial assistance for learners whose task-irrelevant responses to stressful conditions tend to dominate. Further research is required to identify effective remediation procedures which would minimize the debilitating effects of anxiety. Research of this sort requires the manipulation of variables which more directly influence cognitive and affective processes associated with recitation.

Recitation Strategies: The Effect of Rates and Schedules Of Verbal Responses on Retention*

Charles B. Schultz

Recitation strategies have had a long and persistent educational tradition, surviving changes in philosophical orientation, social pressures, and pedagogical attack (Hoetker and Ahlbrand, 1969). While attempts have been made to describe recitation as an instructional strategy characterized by a high rate of student-teacher verbal exchanges, by exchanges of an empirical or factual nature, and by a high ratio of teacher-pupil activity (Bellack, et al., 1966), systematic analyses of the psychological processes underlying recitation are rare. The purpose of this study was to explore the effects of stress produced by recitation strategies on retention for learners who differ in their responses to anxiety.

Recitation is an instructional practice in which learners respond to teacher questions by verbally presenting information that was previously given them. The recitation questions typically are asked in rapid-fire sequences, usually requiring "factual" knowledge or "rote

---

* The cooperation of the administration, faculty, and students of the Bald Eagle Area High School, Wingate, Pennsylvania, in which this experiment was conducted, is gratefully acknowledged. The assistance of Gerald R. Wiser, Principal, William H. Dreibelbis, Guidance Counselor, and John Aliverinini, of the Pennsylvania State University, was particularly appreciated.

memorization." Presumably, the functions of recitation are twofold: Recitation permits the instructor to judge student learning and is therefore evaluative and it strengthens desired learner responses through practice and is therefore instructive. Both functions have implications for the psychological processes assumed to underlie recitation.

An analysis of the instructional function of recitation suggests that there are at least two learning processes involved: direct learning (in which the learner responds to cues and is reinforced) and vicarious learning (in which the learner is presented cues but observes another person respond and receive reinforcement). Both overt rehearsal associated with direct learning and covert rehearsal associated with vicarious learning are assumed to occur during recitation and thereby strengthen the desired response. The effect of rehearsal may be to maintain information in short-term storage, making its coding and subsequent transfer to long-term storage more likely (Atkinson and Shiffrin, 1968). Thus, if the instructional function of recitation is to be effectively implemented, the learner must rehearse and code the desired responses, permitting their transfer to long-term storage.

There are affective as well as cognitive consequences of recitation which appear to be associated with its evaluative function, i.e., with the fact that the responder is judged. Anxiety tends to be generated by the presentation of questions (Kubis, 1948) and the anticipation or fact of speaking before a group (Zajonc, 1966). Thus, there is some support for the intuitive judgment that salient features of the recitation situation are also stressors for the participant. Presumably, the threat posed by these stressors to the learner's self-esteem is heightened when he expects his verbal performance to be evaluated by

teachers and/or peers. The resulting anxiety-produced responses may have been what an early observer of instructional strategies noted when she ascribed "high strung nervous tension" to students in recitation settings (Stevens, 1912, p. 12).

The responses produced by anxiety may be task-relevant in that they facilitate task completion or they may be task-irrelevant in that they interfere with task completion (Mandler and Sarason, 1952). When the relevant responses exceed the amount and strength of the irrelevant responses associated with a particular task such as a test, learners can benefit from anxiety produced by the task (facilitating anxiety). Other students, for whom irrelevant responses dominate in amount or strength, tend to block, freeze, or otherwise are unable to acquire information or produce answers they have acquired (debilitating anxiety).

The present investigation assumed that the stressful conditions of recitation are capable of producing facilitating and debilitating effects on retention similar to those experienced on a test. Task-relevant responses in the recitation situation include rehearsal and coding necessary to transfer information from short- to long-term storage. For recitation, to a greater extent than for many other learning situations, irrelevant responses such as those elicited in anticipation of being called upon and in relief after not being called upon are assumed to compete with rehearsal and coding. Accordingly, the likelihood of information storage for later recall is reduced. In addition, the task-relevant responses associated with recitation may also facilitate the examination and subsequent rehearsal and acquisition of information which is closely related to the instructional topic and perhaps even necessary for its understanding, but which has not been

specifically included in the answers to the recitation questions. The dominance of task-irrelevant responses associated with recitation may inhibit the examination, rehearsal, and acquisition of such information.

The facilitation or inhibition of retention is assumed to be influenced by at least two characteristics of recitation which may regulate the amount of stress recitation produces. Presumably, by increasing the rate at which the learner recites to a point immediately below which habituation occurs, anxiety is heightened. In addition, when the learner does not know which answers he will be required to recite, the instructional situation is less certain and more anxiety producing than when he knows beforehand exactly which answers he will be required to recite.

In summary, when anxiety produced by teacher questions and student recitation before a group results in competing, task-irrelevant responses, rehearsal is impaired and retention reduced; when anxiety facilitates the production of task-relevant responses, rehearsal is enhanced and retention facilitated. The primary hypothesis to be derived from this rationale is: Recitation strategies that produce high anxiety (high rate of response and no knowledge of recitation turn) inhibit retention for learners characterized by debilitating anxiety and improve retention for those learners characterized by facilitating anxiety. A corollary hypothesis is that recitation strategies which are not stressful (low rate of response and knowledge of recitation turn) will have little or no differential effects on the performance of persons characterized by these personality tendencies. These effects are assumed to be reflected in tests of acquisition of material learned directly and vicariously and evidenced on measures of recitation and incidental retention.

Method

Design

The Achievement Anxiety Test (Alpert and Haber, 1960) was administered to a potential pool of Ss three weeks before some were selected to participate in the experiment. The experimental sessions, which were conducted in groups of six Ss, consisted of the following phases: 1) the rapid presentation of 36 slides which described an experiment by Festinger and Carlsmith (1959), 2) the distribution of printed versions of the slides, 3) the recitation of answers to 20 questions which were asked by E and for which answers were readily available in the printed versions of the slides, and 4) the administration of a multiple-choice test of the information contained on the slides.

Stressors to arouse anxiety were manipulated by varying the rate and schedule of recitation. These manipulations, though more precisely defined, were closely analogous to typical classroom procedures. In each group of six Ss, two made seven responses to the 20 recitation questions (35%); two made three responses (15%); and two made no responses. Moreover, half of Ss knew exactly when it was their turn to recite, i.e., the determined schedule, and half did not, i.e., the undetermined schedule. Finally, half of the Ss in each of the above conditions were judged to be facilitators by the Achievement Anxiety Test (AAT) and half were judged to be debilitators. Facilitators were defined as high scorers on the facilitating scale ($\overline{X}$ = 29.42) but low scorers on the debilitating scale ($\overline{X}$ = 20.72) while debilitators were low scorers on the facilitating scale ($\overline{X}$ = 19.33) but high scorers on the debilitating scale ($\overline{X}$ = 34.17). These manipulations imply a

3 x 2 x 2 factorial analysis of variance with three rates of responding (7, 3, and 0), and two schedules of responding (determined and undetermined), and two personality types (facilitators and debilitators).

Subjects

The S pool for the present experiment consisted of 264 eleventh and twelfth grade high school students who had volunteered to take the AAT. Of these, 36 facilitators and 36 debilitators were selected from the academic sections of the school to participate in the experiment on the basis of their AAT scores.

An attempt was made to assign three debilitators and three facilitators to each experimental session; however, because of conflicts encountered in Ss' schedules it was not always possible to maintain this balance. Therefore, additional experimental sessions were conducted which contained "filler" Ss who were selected from the same classes as the regular Ss and whose responses were not included in the analyses. These extra sessions permitted the assignment of the 36 facilitators and 36 debilitators to an equal number of rate and schedule conditions. Each experimental session was randomly designated as a group which would receive a determined or undetermined schedule and the six Ss within each experimental session were randomly assigned to the rate of responding conditions.

Stimulus Materials

The 36 slides used during the first phase of the experiment contained several sentences or a brief paragraph which described an aspect of the assumptions, rationale, procedures, predictions, or results of the Festinger and Carlsmith (1959) study of the cognitive effects of

forced compliance. This topic was selected because it was assumed that the Ss would be unfamiliar with it. The original account was modified to make it easier for high school students to understand. The printed material consisted of reproductions of the 36 slides in booklet form. Several examples of the slides and their corresponding booklet items are presented below.

> Description. The boring task consisted of counting out twelve spools from a large container, placing them on a tray, emptying the tray into a different container, and then refilling it with twelve more spools. This was done for one hour.
>
> Definition. A conflict is created by a situation in which an individual acts inconsistently with his beliefs or attitude. A person will usually try to reduce conflict.

The 20 recitation questions required brief, factual answers which could be obtained easily from the printed versions of the slides. As an illustration, the questions for the above items were: What was the boring task and how long was it performed? and What is the psychological definition of conflict? In order to control for possible differences in question difficulty, the random assignment of questions to the three and seven rate of response conditions was recycled for each set of experimental sessions. A set consisted of a determined schedule group and an undetermined schedule group for which the assignment of questions to the response rate conditions was matched. Since it was necessary to inform the determined schedule Ss of the questions they were to answer, a red number indicating the location of the answer to the recitation question each S would be called upon to answer and its place in the 20 question sequence was placed next to the appropriate item in each booklet. For the undetermined schedule, all 20 recitation items were numbered in red. Thus, the Ss in the response rate conditions of the

two groups which comprised a set of experimental sessions (i.e., a determined and an undetermined schedule group) answered the same questions but used different sets of booklets, each numbered in red according to the schedule for which it was used.

Procedures

After entering the experimental room, the six Ss were seated in a small semi-circle, facing E. Thus, each S could easily be seen as he recited by the other Ss in the experimental session and by E. The Ss were informed that the purpose of the experiment was to examine the effectiveness of several teaching methods and the major phases of the "experimental lesson" were described. For the first phase, the slides were presented at a rapid, eight second pace, allowing Ss only enough time to scan their content. Before the recitation phase, Ss were instructed to repeat the answers verbatim and to listen closely to the answers others gave because "they would be tested on all the material presented during the lesson." At this point, Ss in the three rate of responding conditions were informed of the number of questions they would be required to answer. In addition, the determined schedule condition received the following instructions:

> You will know exactly which questions you will have to answer. The numbers written in red next to some of the paragraphs indicate which of the twenty questions you will be called upon to answer in class. The correct answer can be found in the paragraph next to the red, handwritten number. For example, if '3' were written in red next to one of the paragraphs, it would mean that you will be asked the third recitation question and that the answer is in the paragraph next to the number '3'.

Instructions for the undetermined schedule were as follows:

> You will not know when you will be called upon to answer these questions. You could be called upon at any time. The correct answer to the questions can be found in the paragraph next to the red handwritten numbers. For example, if '3' were written next to one of the paragraphs, it would mean the answer to the third recitation question could be found in that paragraph.

The recitation questions themselves were asked in a matter-of-fact tone. When an incorrect answer was given, E asked S to look at the printed material again and find a different answer. After the recitation period, a multiple-choice test was administered along with a post-experimental questionnaire. The 30 item test contained 20 items which required information rehearsed during the recitation period (i.e., recitation retention) and ten items which required information included in the printed material and relevant to the experimental topic, but which Ss were not directed to recite and thereby rehearse during the recitation session (i.e., incidental retention). The post-experimental questionnaire included the following question: How much tension did you feel during the question and answer period? The Ss rated their tension on a five point scale labeled "no tension" at one extreme, "moderate tension" at the midpoint, and "very tense" at the other extreme. Finally, Ss were asked if they had prior knowledge of the experimental topic.

Results

None of the Ss responded positively to the question of whether they had prior knowledge of the experimental topic nor could they offer adequate definitions of cognitive dissonance. Accordingly, differences in retention scores cannot reasonably be attributed to differences in prior knowledge.

The three-factor analysis of variance of reported tension was made to determine the extent to which anxiety was induced. Since this analysis yielded negligible differences in tension between the two schedules ($F < 1$), the three-dimensional design was collapsed across schedules to examine the effects of the remaining factors, (i.e., rates of responding and personality differences) on reported tension. The means and standard deviations for this analysis are displayed in Table 1.

Regardless of schedule, recitation appears to have produced stress which was most evident at moderate levels of verbal responding. The analysis of the effect of response rate on tension yielded F (2,66) = 4.62, $p < .01$. Pairwise comparison of the means, using multiple t ratios, indicated that less tension was reported in the no response condition than in either the seven response (one-tailed t (66) = 1.88, $p < .05$) or the three response conditions (t (66) = 3.01, $p < .01$).

The same analysis showed that the effect of personality on tension was significant F (1,66) = 8.80, $p < .004$, indicating that debilitators ($\bar{X}$ = 3.97) reported more tension than facilitators ($\bar{X}$ = 2.99). The analysis also yielded a significant interaction between personality and response rate, F (2,66) = 3.40, $p < .04$, in which no difference between facilitators and debilitators was obtained in reported tension in the no response condition while debilitators reported more tension than facilitators in both the three response condition (t (66) = 2.30, $p < .05$), and seven response condition (t (66) = 3.19, $p < .01$). Thus, differences in tension reported by facilitators and debilitators increased as response rate increased. This trend revealed a tendency for debilitators to experience considerable tension at either the three or seven response rates and a tendency of facilitators to experience little tension at the seven response rate.

Table 1

Means and Standard Deviations of Reported Tension for Personality Type and Response Rate

Personality Type		Rate of Response		
		0	3	7
Facilitators	$\overline{X}$	2.12	2.50	1.92
	SD	1.11	1.13	.87
Debilitators	$\overline{X}$	1.96	3.58	3.42
	SD	1.21	1.24	1.29

A three-factor analysis of variance was made of the effects of rates of responding, schedules of responding, and personality differences on both recitation retention and incidental retention. It was expected that the undetermined schedule, which was assumed to be more anxiety-producing, would result in greater differences between facilitators and debilitators across the rates of responding conditions than the determined schedule. The analysis of the effect of schedules of responding on recitation retention yielded $\underline{F}$ (2,60) = 6.83, $\underline{p} < .01$ implying that the undetermined schedule ($\overline{X}$ = 11.44) resulted in the acquisition of more information than the determined schedule ($\overline{X}$ = 9.64). However, interactions were not obtained between schedules of responding and any of the other independent variables ($\underline{F} < 1$) on either measures of recitation or incidental retention. Therefore, the original three dimensional design was collapsed across schedules to examine the effects of rates of responding and personality differences for the two dependent variables. The means and standard deviations for this analysis are displayed in Table 2.

The analysis of variance of the effects of personality on recitation retention yielded $\underline{F}$ (1,66) = 9.29, $\underline{p} < .003$, in which facilitators ($\overline{X}$ = 11.61) acquired more information than debilitators ($\overline{X}$ = 9.47). Differences among the rate of responding conditions were also significant, $\underline{F}$ (2,66) = 3.61, $\underline{p} < .03$. Pairwise comparisons among the means indicated that the seven response condition did not result in as much retention as the three response ($\underline{t}$ (66) = 2.38, $\underline{p} < .05$) or no response conditions ($\underline{t}$ (66) = 2.27, $\underline{p} < .05$). A major interest in the data obtained from the recitation retention measure regards the hypothesized interaction in which differences in retention between facilitators and debilitators

Table 2

Means and Standard Deviations of Recitation Retention (RR) and Incidental Retention (IR) for Personality Type and Response Rate

Measure	Personality Type		Rate of Response 0	3	7
RR	Facilitators	$\bar{X}$	11.91	12.58	10.33
		SD	3.15	2.07	2.57
	Debilitators	$\bar{X}$	10.42	9.92	8.08
		SD	4.84	2.75	2.19
IR	Facilitators	$\bar{X}$	4.42	5.00	5.92
		SD	1.88	1.60	1.62
	Debilitators	$\bar{X}$	4.17	3.58	4.17
		SD	1.53	.90	1.85

were expected to increase across increasing response rates. Although the response rate by personality interaction was not significant ($F < 1$), the recitation retention scores for debilitators were in a direction which was consistent with the hypothesis. The decrease in scores across increasing response rates is most evident in the analysis of the differences between the seven ($\bar{X} = 8.08$) and no response ($\bar{X} = 10.42$) condition which yielded t (66) = 2.01, $p < .05$. Contrary to expectations, the retention scores of facilitators in the seven response condition ($\bar{X} = 10.33$) were lower than those for the three response condition ($\bar{X} = 12.58$), although this difference did not reach traditional levels of significance ($p < .05$).

The analysis of variance of the effects of rates of response and personality differences were made for incidental retention. The means and standard deviations for this analysis are displayed in Table 2. The analysis of the effect of personality on incidental retention was significant, F (1,66) = 9.17, $p < .004$, indicating that facilitators ($\bar{X} = 5.11$) retained more information than debilitators ($\bar{X} = 3.97$). However, the effects of response rate on incidental retention and the response rate by personality interaction were not significant ($F < 1$). The incidental retention scores of facilitators are of particular interest in regard to the hypothesis of the present experiment because their scores increased as the response rate increased. Pairwise comparisons of the differences between the no response ($\bar{X} = 4.42$) and seven response conditions ($\bar{X} = 5.92$) yielded t (66) = 2.31, $p < .05$.

In summary, the hypothesized interaction between levels of experimentally induced anxiety and the recitation and incidental retention of facilitators was not obtained. However, the recitation and incidental

retention of facilitators was superior to that of debilitators, suggesting that recitation strategies differentially affect the amount of information retained by these two personality types. This difference was most apparent in the recitation retention scores of debilitators which decreased across increasing rates of response and in the incidental retention scores of facilitators which tended to increase across increasing rates of response.

The response rate manipulations were attempts to induce different levels of anxiety. Since an interaction was not obtained between response rate and personality on measures of retention, a *post hoc* analysis was conducted to determine whether a direct relationship existed between the tension *Ss* reported and recitation retention for facilitators and a negative relationship for debilitators. According to this analysis, no relationship was found between reported tension and recitation retention for facilitators ($r = .02$) while a slight, negative relationship ($r = -.14$, $p > .05$) was obtained for debilitators.

Discussion

According to the findings obtained in the present study, in recitation settings facilitators retained more of the information they were directed to examine than debilitators. Moreover, the differences between facilitators and debilitators were maintained for retention of information closely related to the topic and which *Ss* had an opportunity to examine, but were not directed to do so. These differences were consistent with the interpretation that anxiety produces responses which facilitate task completion for some individuals and which interfere with task completion for others. The retention scores for debilitators,

which tended to decrease as experimentally induced anxiety increased, were particularily supportive of the present hypothesis and of the competing response interpretation (Mandler and Sarason, 1952). In addition to lowering retention scores, the generation of competing, task-irrelevant responses by debilitators may account for the high arousal reflected in the tension they reported when required to recite. This interpretation is consistent with the findings that debilitators who recited reported more tension than either facilitators or debilitators who did not recite and suggests that recitation produces relatively intensive response competition in debilitators.

When responses which compete with those required for task completion are conceived of as intrusions of irrelevant thoughts (Sarason, et al., 1960), the responses of debilitators to anxiety assume the form of informational inputs which tend to "overload" short-term storage. The effect of such informational inputs is to limit the capacity of short-term storage to hold task relevant-information, causing the decay of relevant information and thereby prohibiting its transfer to long-term storage. In recitation settings these intrusions, which debilitators appear to enter into short-term storage more often than facilitators, may include thoughts associated with anticipation of reciting and relief when not called upon. The overload of short-term storage is one explanation of recent findings by Siever, Kameya, and Paulson (1970). These investigators found that without memory supports, problem-solving of high-anxious Ss was poorer than that of low-anxious Ss; however, provision of memory supports facilitated problem-solving for highly anxious individuals to a point where the two groups did not differ.

Memory supports were assumed to supplement the function of short-term storage, thereby overcoming the disruption caused by the intrusions of competing responses. Thus, both the Sieber study and the present experiment suggest that anxiety can have disruptive effects on short-term storage.

The failure to obtain the hypothesized interaction between personality and experimentally induced anxiety on measures of recitation retention can be attributed, in part, to the relatively poor performance of facilitators in the seven response condition. The low retention scores of these Ss may have been due to their habituation to the effects of questions and recitation with the consequent reduction of anxiety (Kubis, 1948). Habituation is suggested by the low ratings of tension reported by facilitators in the seven response condition. In this regard, it appears that the habituation to stressful conditions by facilitators is more rapid than that of debilitators.

The undetermined schedule resulted in greater recitation retention than the determined schedule for both facilitators and debilitators. The superiority of the undetermined schedule suggests that certain characteristics of that schedule produced overriding effects which did not contribute to the expected interaction. For example, the undetermined schedule may have demanded more intensive examination of the experimental materials than the determined schedule. Since Ss in the undetermined schedule could be called upon to recite at any time, it is likely that they searched for more answers and covertly rehearsed them, resulting in greater information storage, than Ss who knew which answers they would recite. In this regard, differences between the schedules were not evident in scores for incidental retention, which measured

acquisition of information for which Ss did not have to be prepared to recite, and therefore dod not require differential search and rehearsal activities.

While instructional implications can be drawn from this experiment including the use of undetermined schedules and moderate rates of responding, neither suggestion is likely to improve the relatively poor performance of debilitators in stress-producing conditions like recitation. The development of effective remediational techniques requires further experimentation. In the present study, the variables which were manipulated were the most apparent elements of the recitation setting. Remediational techniques may be identified by the manipulation of variables which are more directly associated with the cognitive and affective processes assumed to occur during recitation. The former processes suggest manipulations which would require debilitators to code information (e.g., instructions to "translate" answers rather than verbatim repetitions) or the provision of labels or "anchors" (e.g., the use of advanced organizers) which may also facilitate coding and the consequent transfer of information to long-term storage. The affective processes suggest manipulations which would vary the threat to S's self-esteem which is apparently posed by the recitation strategies.

References

Alpert, F. and Haber, R. N. Anxiety in academic achievement situations. Journal of Abnormal and Social Psychology, 1960, 61, 207-215.

Atkinson, R. C. and Shiffrin, R. M. Human memory: A proposed system and its control processes. In Spence, K. W. and Spence, J. T. (Eds.), The psychology of learning and motivation: Advances in research and theory. Vol. 2. New York: Academic Press, 1968.

Bellack, A. A., Kliebard, H. M., Hyman, R. T., and Smith, F. L. The language of the classroom. New York: Teachers College Press, 1966.

Festinger, L. and Carlsmith, J. M. Cognitive consequences of forced compliance. Journal of Abnormal and Social Psychology, 1959, 58, 203-210.

Hoetker, J. and Ahlbrand, W. P. The persistence of recitation. American Educational Research Journal, 1969, 6, 145-167.

Kubis, I. F. Adaptation of the psychogalvanic response to a visual, auditory, and ideational stimulus. American Psychologist, 1948, 3, 256.

Mandler, G. and Sarason, S. B. A study of anxiety and learning. Journal of Abnormal and Social Psychology, 1952, 49, 166-173.

Sarason, S. B., Davidson, K. S., Lighthall, F. F., Waite, R. R., and Ruebush, B. K. Anxiety in elementary school children. New York: Wiley, 1960.

Sieber, J. E., Kameya, L. I., and Paulson, F. L. The effect of memory support on the problem-solving ability of test-anxious children. Journal of Educational Psychology, 1970, 61, 159-168.

Stevens, R. The question as a measure of efficiency in instruction: A critical study of classroom practice. New York: Teachers College, Columbia University. Contributions to Education No. 48, 1912.

Zajonc, R. B. Social psychology: An experimental appraoch. Belmont, California: Wadsworth, 1966.

Summary

Satiation of Divergent and Convergent Thinking and Its Effect on the Need for Novelty*

Study Director: John R. Silvestro

Advisor: Francis J. Di Vesta

Technical Problem

This study investigated the effects of satiating learners with either divergent or convergent thinking activities on their inferred desire to seek out novel stimuli. Research by Houston and Mednick (1963) had shown that the need for novelty was significantly greater among high-creative Ss than among low-creative Ss. The present author reasoned that it was not simply the level of creativity which Ss brought into the experiment that determined how strong their desire for novelty would be. Rather, it was assumed that certain specific antecedent conditions peculiar to each S, prior to his entry into the experiment, differentially affected the strength of the measured need for novelty. Also, it was predicted that differences between high- and low-creatives as determined by the Remote Associates Test (RAT) would be minimal.

---

* This study was conducted as a Master's thesis in the Department of Educational Psychology and was supported, in part, under the present contract. The complete thesis will be published as a Technical Report (Technical Report No. 1).

General Methodology

An experiment was conducted in which high- and low-creative Ss were divided into divergent satiation and convergent satiation treatment groups. The divergent satiation groups were given a series of creative, imaginative, and flexible tasks, while the convergent satiation groups were given a series of highly structured, common, simple tasks, that called for one and only one appropriate response. Following the satiation condition, each S's inferred need for novelty was measured. Ss were shown 180 slides. On each of 160 slides was a pair of words, a noun and a non-noun. For half of the high- and low-creatives, when a noun was selected, the E responded verbally with a novel association of that noun. When a non-noun was selected by S, the E responded with a common association. For the other half of the Ss the procedure was reversed, with nouns eliciting common responses, and non-nouns eliciting novel responses. The remaining 20 slides were filler items used to prevent S from gaining insight into E's intent.

Technical Results

The results of this experiment indicated that Ss satiated with convergent thinking tasks displayed a stronger need for novelty than Ss satiated with divergent thinking tasks. These results were obtained regardless of the level of creativity of S, although the differences tended to be slightly greater between the groups of high-creative Ss than they were between the groups of low-creative Ss. These differences were attributed to the possible tendency of low-creatives to have a somewhat lower need for novelty. The effect of type of word class used for reinforcement was significant among high-creatives but not for

low-creatives; high-creative Ss selected more nouns associated with novel responses than they did non-nouns. The optimal condition for arousing a preference for novel associations was convergent satiation in combination with nouns eliciting the novel responses. The combination which produced the least influence on the receptivity to novel associations as reinforcing stimuli was satiation on divergent thinking tasks with non-nouns eliciting the novel responses.

Educational Implications

The basic implications from this study for use in instructional settings are twofold: First, a need for novelty can be aroused in any learner, whether a high- or low-creative person, by adequate control of the antecedent conditions. Conversely, the need for novelty may be hampered or the need for normative behavior encouraged by over-exercise on tasks requiring creative effort. Secondly, it is inferred from this study that moderate emphasis on creativity ought to be adopted by instructors since overuse may lead to satiation of creative stimulation and thereby decrease the need for novelty along with its desirable effects on performance and learning.

Implications for Further Research

In order to render the results more generalizable it is recommended that more research be undertaken to delineate the quantitative (levels) and qualitative characteristics of divergent thinking that serve to sustain the need for novelty for optimal periods of time. It would appear that a quasi-naturalistic study comparing the subsequent preference for novel stimulation of students who are taught by a convergently oriented teacher with students taught by a divergently oriented teacher

would provide useful supplementary data to that obtained in the present study. Such research could be instrumental in ascertaining the interaction of convergent or divergent satiation with high and low IQ or other aptitudes.

References

Houston, J. & Mednick S. Creativity and the need for novelty. Journal of Abnormal and Social Psychology, 1963, 66, 137-141.

Summary

Small Group Verbal Presentations, Anxiety Level, and Learning in Instructional Settings

Paul Weener

Technical Problem

This study investigated the effects on learning of student verbal presentations and the interaction of these effects with learner anxiety levels. It was hypothesized that the effects of student verbal presentations can be explained in terms of the arousal effects resulting from this activity. According to this explanation, optimal learning occurs if the level of arousal is low during the early stages of the learning of new material and increases as the material becomes better learned.

General Methodology

An experiment was conducted in which four treatment groups were formed in terms of the Ss activities during two study periods which followed the viewing of an instructional film. The four groups respectively engaged in the following sequence of study activities: (1) study alone - study alone, (2) verbal presentation - study alone, (3) study alone - verbal presentation, and (4) verbal presentation - verbal presentation.

Technical Results

Recall scores were obtained on an essay and an objective test immediately following the study activity and again one week later. No significant mean differences among the groups were obtained on either the immediate posttest or the delayed retest, but the trend of the results favored the groups which participated in verbal presentation activities during the study period. The two groups which participated in verbal presentation activities during the second study period performed consistently higher than the two groups which studied alone during the period immediately preceding the test. Although a measure of Debilitating Anxiety consistently correlated negatively with performance and a measure of Facilitating Anxiety consistently correlated positively with performance, there was no pattern of correlations across the four treatment groups which indicated that the treatment conditions interacted with individual differences in anxiety levels.

Educational Implications

Very little can be said about the implications of chance level differences in an experiment. If the trends observed in this study could be amplified in subsequent studies, a case could be made for advocating the use of verbal presentation techniques during the later stages of a period of study.

Implications for Future Research

The experimental treatments should be applied over a longer period of time in order to amplify differential treatment effects if present.

Small Group Verbal Presentations, Anxiety Level, and Learning in Instructional Settings

Paul Weener

At least three different theoretical rationales can be used to explain the effects on learning of student verbal presentations in the classroom. The effects could be explained in terms of (1) the active associational and coding processes which are induced by verbal presentation instructions, (2) the overt nature of the response, and (3) the arousal resulting from verbal presentation.

The first explanation would argue that verbal presentation instructions would result in active associational and coding processes which would facilitate learning. These are the processes of "making meaningful," of "putting into one's own words," of "assimilating into cognitive structure." These processes include the substitution of familiar words and phrases for unfamiliar words and phrases, the application of some meaningful mnemonic to remember the overall structure of the presentatation and the transformation of the stimulus material into a set of symbols which can be stored and processed effectively. If classroom verbal presentation conditions produce these active coding processes, then one would argue that student verbal presentations should facilitate learning - and particularly long term learning - as compared to an instructional setting in which the student did not have the opportunity to verbally present.

Verbal presentation could also have facilitative effects as compared to more passive study conditions if one maintains that overt responses are remembered better and longer than covert responses. There have been conflicting research findings regarding the role of verbalization in learning (Cofer, 1960), but it can be argued from a motor theory of memory that an overt response, because it involves muscle responses in addition to the covert mental responses should be better remembered than simply the covert response.

If the effects of verbal presentations in instructional settings is primarily in terms of arousal level, then theories of the effects of arousal (e.g., Spence & Spence, 1966), can be called on to explain the effects of student verbal presentations. Just as Zajonc (1966) used arousal concepts to explain the effects of working in groups as compared to working alone, the effects of verbal presentation conditions can be explained in terms of the concomitant levels of arousal produced by such conditions. According to the arousal interpretation, the effects of making a verbal presentation in the presence of others tends to increase a person's arousal level which in turn increases the probability of a dominant response being emitted. During the early stages of the learning, wrong responses tend to be dominant, and the frequency of the wrong responses would be increased when the arousal level is raised by the verbal presentation conditions. After the correct responses have become dominant in the later stages of learning, the verbal presentation in the presence of others would facilitate learning because the accompanying heightened level of arousal would increase the emission of dominant responses which now contain more correct responses. Zajonc concluded his interpretation of the arousal effects of working in the

presence of others by stating - somewhat facetiously - that students should study quietly alone until they have learned the materials well and then take an exam over the material on a stage in front of an audience of people. The present study was developed with a similar interpretation of the effects of verbal presentation.

The hypotheses in this study are based on a model which predicates that the functional level of anxiety in a learning setting is the product of the arousal invoking characteristics of the instructional setting and the individual's susceptibility to the arousal-invoking characteristics of the instructional setting. The functional level of anxiety is then curvilinearly related to performance, with combinations of low anxiety situational characteristics and low susceptibility as well as high anxiety situational characteristics and high susceptibility resulting in performance which is lower than the performance resulting from a more moderate level of functional anxiety.

The first hypothesis in this study states that the most effective sequence of events in an instructional setting, following the presentation of new material, is (1) study alone, (2) present verbally to peer group, and (3) take test on material. The Zajonc interpretation implies that in order to produce optimal learning the level of arousal should increase as the stimulus materials become better learned. That is, as the strength of the correct response tendencies increase, the level of arousal required for optimal performance would also increase.

The second hypothesis states that arousal, induced by verbal presentation, is mediated by the student's susceptibility to the effects of achievement related anxiety. Since performance is curvilinearily related to levels of arousal, it is assumed that for students with low

susceptibility to arousal the effects of making a verbal presentation would result in an optimal level of arousal, whereas for a student with a high susceptibility to arousal the effects would result in a debilitatingly high level of arousal.

Method

Design

Four treatment conditions were defined in terms of the Ss activities during two study periods which followed the presentation of the instructional materials. The four groups and their activities were as follows:

	Activity	
Group	First Study Period	Second Study Period
1	Study Alone	Study Alone
2	Verbal Presentation	Study Alone
3	Study Alone	Verbal Presentation
4	Verbal Presentation	Verbal Presentation

The dependent measure was performance on an objective test and an essay test. Three repeated measures on each of the two types of tests were obtained, resulting in a 4 x 2 x 3 design with repeated measures over the last two factors. The repeated measures were an immediate posttest, a delayed retest, and a delayed parallel form retest.

Subjects

Ninety-four Pennsylvania State University undergraduate students from the introductory educational psychology course participated as Ss in the experiment. Ss volunteered to participate and were awarded points toward their course grades for participation. All Ss had

participated in a group testing session two weeks prior to the beginning of the experiment. During this session several individual difference measures were obtained for use in later research.

Materials

A 22 minute movie, titled The Tibetan Traders, served as the stimulus presentation in the experiment. The film told the story of a village of traders who moved with the seasons in order to trade for the variety of goods needed for survival.

Two parallel forms of a test were developed to measure knowledge of the material presented in the film. Each test had eleven multiple-choice questions and seven short answer questions together comprising the objective part of the test, in addition to one general essay question.

The State Anxiety Inventory (STAI) was used to measure the levels of anxiety produced by each of the four experimental conditions. This instrument was developed by Spielberger, Forsuch, & Lushene (1968), for the purpose of measuring personal anxiety level in specific situations. The scale, entitled, "Self-Analysis Questionnaire" consists of thirteen items, e.g., "I was calm," "I was tense." The S responds to each item with one of the following alternatives: (1) not at all; (2) somewhat; (3) moderately so; (4) very much so. A simple summation with reversal of negatively worded items yields the total score. Spielberger (1968, p. 11) reported internal consistency reliabilities of .88 and .90 for male and female undergraduates respectively.

The Achievement Anxiety Test (AAT), developed by Alpert & Haber (1960), consists of a total of nineteen items. The items are divided

into a Facilitating and a Debilitating Anxiety Scale which measure the extent to which achievement - related pressures are reported to facilitate or debilitate academic performance respectively.

Procedure

Ss met in groups of six for the showing of the film and were assigned at random to one of the four treatment conditions. They were seated in two rows of chairs and given a tablet and pencil. Instructions were then read to each group by the E. In all treatments, Ss were told they could take notes and that a test would be given following the experiment. In treatment conditions 2, 3, and 4, the groups were divided into 2 three-person groups. Within each of these subgroups, the Ss were assigned a letter A, B, or C.

Treatment 1 groups were instructed that a film would be shown followed by an eighteen-minute study period during which they would study by themselves. Treatment 2 groups were told that a film would be shown followed by two nine-minute study periods. During the first study period, each person within the subgroup would make a 3 minute presentation summarizing the film, in the order A-B-C. During the second study period, they would study alone the notes they had taken. Treatment 3 instructions were the same as treatment 2 instructions, except that the "study alone" period preceded the "verbal presentation" period. Treatment 4 groups were instructed that during each of the study periods, each person would make a three minute presentation to their subgroup.

The E left the experimental room after the instructions had been read, and monitored the rest of the experimental period from a small

observation room behind the experimental room. The film was projected through the one-way mirror and timing instructions were given from the observation room.

After the final study period ended, copies of the test were handed out. The Ss were given ten minutes for the objective part of the test and seven minutes for the essay question. Three of the Ss in each group were given one form of the test and three were given a parallel form of the test.

Following the test, the STAI was administered with an average administration time of about three minutes. All Ss were then reminded to return one week later for the second part of the experiment, and that they would not obtain any credits for their participation unless they returned at this time. When Ss returned one week later, they were administered the same test they had taken a week earlier, as well as the parallel form of the test.

Twenty-four Ss were run in treatments 2, 3, and 4. One of the sessions in treatment 1 had only 4 Ss participating, leaving a total of 22 Ss in group 1. Two Ss - one in group 1 and one in group 3 - did not return for the retesting session.

Scoring

The seventeen objective test items were scored right or wrong. The essay question was scored by assigning one point to each independent and dependent clause which was factually correct. Redundant statements, incorrect statements, statements not based on the movie, and irrelevant statements were not given any points.

Results

Six scores are available for each S who participated in both parts of the experiment. An objective test score and an essay test score were obtained for the immediate posttest, the delayed retest, and the delayed parallel form retest. The means and standard deviations for the six scores for each experimental group are presented in Table 1.

The first hypothesis predicted that Group 3 would perform better on the tests than any other group. Two 4 x 3 analyses of variance were carried out, one on the objective test scores and another on the essay test scores to test this hypothesis. The four experimental groups comprised the levels of the within-subject factor. The analysis on the objective test data indicated no significant differences among experimental groups, F (3,88) = .58, $p > .05$. The three test scores were significantly different, F (2,176) = 44.3, $p < .01$, but the interaction between experimental group and test was not significant, F (6,176) = 1.18, $p > .05$.

The analysis on the essay test also indicated no significant differences among experimental groups, F (3.88) = 1.15, $p > .05$. The three test scores were significantly different, F (2,176) = 11.73, $p < .01$, but the interaction between the two factors was not significant, F (6,176) = .30, $p > .05$.

A Newman-Keuls analysis of the differences among the three objective test means indicated that the immediate posttest mean score was not significantly different from the delayed retest mean, but that both the immediate posttest and the delayed retest were significantly different from the parallel form retest ($p < .05$). The Newman-Keuls analysis of

Table 1

Means and Standard Deviations of Six Test Scores

for Four Experimental Groups

Test Scores			Groups One*	Two	Three	Four
Immediate Posttest	Objective	$\bar{X}$	12.86	12.96	13.74	13.54
		S	3.23	2.46	1.84	2.40
	Essay	$\bar{X}$	5.00	5.71	5.91	6.08
		S	1.18	2.37	2.02	2.19
Delayed Retest	Objective	$\bar{X}$	12.38	12.29	12.09	12.42
		S	2.48	2.53	1.76	2.52
	Essay	$\bar{X}$	4.43	4.38	5.13	5.08
		S	2.20	1.66	2.16	2.50
Delayed Parallel Form Retest	Objective	$\bar{X}$	10.14	10.88	10.65	11.67
		S	2.63	2.13	2.60	2.28
	Essay	$\bar{X}$	4.10	4.87	5.00	4.88
		S	1.81	2.67	2.89	1.80

*Group 1 - Study alone-Study alone

Group 2 - Presentation-Study alone

Group 3 - Study alone-Presentation

Group 4 - Presentation-Presentation

differences among the three essay test means indicated that the immediate posttest was significantly different from both the delayed retest and the delayed parallel form retest but that the other comparisons were not significant.

Hypothesis two stated that the relationship between test performance and debilitating anxiety should be more negative under instructional conditions which induce anxiety than under conditions which result in a more moderate level of anxiety. Specifically, support for this hypothesis would be obtained if the correlation between the Debilitating Anxiety score and test performance were more negative in the group 4, verbal presentation, condition than in the group 1, study alone, condition. Table 2 presents the correlations between the six test scores and the Debilitating and Facilitating Anxiety score for each treatment group.

Hypothesis two was not supported. The differences between the correlations in group 1 and the correlations in group 4 are not significant. There is no discernible pattern across groups within the correlation coefficients obtained with the Facilitating or Debilitating scores. Comparing the relationship between the anxiety measures and performance, 23 out of the 24 correlations of test scores with Debilitating Anxiety were negative, and 23 out of the 24 correlations of test scores with Facilitating Anxiety were positive.

In order to determine the effect of the treatment condition on the S's feelings of anxiety within the experimental condition, an analysis was performed on the STAI scores obtained immediately following the experiment. The means and standard deviations for the four groups were 19.6 (4.1), 19.9 (5.1), 20.7 (7.2), and 22.1 (7.0), respectively. A one-way analysis of variance yielded F (3.90) = .84, $p > .05$.

Table 2

Correlations Between Six Test Scores

and Facilitating and Debilitating Anxiety Scores

		Anxiety Measure							
		Facilitating				Debilitating			
	Group	1*	2	3	4	1	2	3	4
Test Score									
Immediate Posttest	Objective	.61	.02	.00	.37	-.25	-.25	.05	-.10
	Essay	-.05	.17	.37	.13	-.10	-.08	-.53	-.28
Delayed Retest	Objective	.56	.25	.22	.33	-.18	-.35	-.06	-.11
	Essay	.34	.25	.00	.41	-.42	-.26	-.27	-.10
Delayed Parallel Form Retest	Objective	.41	.43	.11	.11	-.36	-.29	-.07	-.30
	Essay	.10	.23	.34	.09	-.09	-.29	-.48	-.07

*Group 1 - Study-Study

Group 2 - Verbal Presentation-Study

Group 3 - Study-Verbal Presentation

Group 4 - Verbal Presentation-Verbal Presentation

Discussion

Statistical support for the hypotheses regarding the effects of student verbal presentations on learning was not obtained. The trends in the data favor the groups in which verbal presentations were part of the study procedure. The "study only" group had the lowest group score on four of the six tests. Either group 3 (study alone - presentation) or group 4 (presentation - presentation) obtained the highest mean score on each of the six tests. The means of the total scores obtained by summing across all six tests are 81.5, 84.0, 87.1, and 87.4 for the four groups respectively. However, the within-group variance is so large that these differences cannot be attributed to the experimental treatments.

The correlational data yielded no statistical support for the hypothesized interaction between treatment condition and level of debilitating anxiety. The only clear indication obtained from this analysis is that Ss who scored higher on the facilitating anxiety scale also scored higher on the tests, and that Ss who scored higher on the debilitating anxiety scale scored lower on the tests. From Table 2, the median value for the 24 coefficients obtained between facilitating anxiety and test scores was +.25; the median value for the 24 coefficients obtained between debilitating anxiety and test scores was -.25.

Tne analysis of situational anxiety scores obtained on the STAI indicated a trend with higher anxiety scores associated with the conditions in which verbal presentations were required. The within-group variance was again too large to conclude that the differences resulted from the experimental treatment.

Three possible explanations can be given to the findings: (1) the theory relating performance on paper-and-pencil memory tests to arousal level is not correct, (2) the experimental treatments were not strong enough to produce arousal differences and performance differences, and (3) the measuring instruments were not sensitive enough to detect real differences. The analysis on the STAI scores gives direct evidence that the second explanation has some credibility. With regard to the measuring instruments, the test-retest reliability of the objective test was .81, no ceiling effect seemed to be operating, but the range of obtained scores was quite small which may reflect the insensitivity of the instrument.

Further research should be done in which the experimental treatment is applied over a longer period of time under carefully controlled, but non-artificial, situations. If the small differences produced in the present research would then increase, an interpretation of the effects of overt verbal presentations in instructional settings could be obtained.

References

Alpert, R. & Haber, R. N. Anxiety in academic achievement situations. Journal of Abnormal and Social Psychology, 1960, 61, 207-215.

Cofer, C. Experimental studies of the role of verbal processes in concept formation and problem solving. Annals of the New York Academy of Sciences, 91, 1960, pp. 94-107.

Spence, J. & Spence, K. The motivational components of manifest anxiety: Drive and drive stimuli. In C. Speilberger (Ed.), Anxiety and behavior. New York: Academic Press, 1966, 291-326.

Spielberger, C. D., Gorsuch R. L., Lushene, R. E. The state-trait anxiety inventory. Tallahassee, Fla.: Psychology Department, Florida State University, 1968.

Zajonc, R. B. Social facilitation. Science, July 16, 1965, 269-274.

PROGRESS REPORTS

Project Ikon: Studies of Imagery and Learning

Study Director: Francis J. Di Vesta

Assistants: Gail Susan Gray, Gary Ingersoll, Edwin Marlow, Steven Ross, Phyllis Sunshine

Purpose

The program described in this progress report is based on the assumption that some adults continue, for whatever reason, a pre-verbal dependence on concrete images in their thinking habits; others, with the development of language, discard images in favor of verbal or other symbolic representations as characteristic modes of thought. It is the purpose of this program to investigate the effects on learning of individual differences in imagery habits as they interact with materials presented by visual and symbolic modes. We expect to investigate both the conditions under which imagery facilitates, and under which it interferes with, performance.

These studies fall neatly into place when viewed from the standpoint of the instructional model presented at the beginning of this report. Within the overall theory, imagery may be considered either an individual difference variable or as a process. As an individual difference variable, imagery functions in that part of the "filter" system which was labeled "modality preference." There, along with modality preferences and other individual difference variables, imagery influences that feature of behavior that has come to be known as "selective attention"

(Broadbent, 1958). As a process, imagery affects the way learning materials are transformed and, accordingly, the rapidity with which they are acquired, the length of time they are stored, the way they are recalled, and the conditions for optimal recall. The importance of imagery in children's behavior has received considerable attention of late. Some of the general notions associated with related theories, however, seem to be equally relevant to information processing in the adult's learning. Thus, Paivio (1970) suggests that the image serves as a "conceptual peg" for storage and retrieval of the response item. Rohwer (1970) concluded that imagery is most effective when a verbal tag is stored with the image (also see the study by Di Vesta and Rickards described earlier in this report.).

There are many notions about imagery, not always consistent with one another, which bear on the present research. For examples: Galton, with whom the first research in this area is typically associated, concluded that an "over-ready perception of sharp mental pictures is antagonistic to the acquisition of habits of highly generalized and abstract thought." Roger Brown states that images cannot comprise the non-linguistic meaning categories that are referents for words. Children depend on straightforward images which are dropped in preference to abstract thinking; they move from concrete perceptual bases to functional bases of classification. Anne Roe, in a study of imagery in scientists found may who used visual imagery in solution of problems.

There are questions raised from a number of quarters regarding the nature of an image.Though it is recognized as a representation in memory of an absent object (or idea) much more needs to be known about its characteristics. Among questions that might be asked are the following:

Is it a picture, abstraction, or template? Does it become altered? Is it elaborated, refined by sharpening, or dulled by leveling over time? If this representation is a commonality of a class of events does such representation continue to exist at highly abstract stages of thought?

Whether one refers to imagery, hypothesis testing, TOTE (Test-Operate-Test-Exit) strategies, verbal mediation, or conceptualization it is evident that there is something going on in the student which is somehow facilitated by concreteness, tags or labels, pictures, instructions, and the like. An understanding of these processes in the adult's learning is intended through the present studies on the functioning of imagery in thought.

Relevance for Instruction

Our progress in research in this area to date attests to the validity of the assumption that reliable individual differences exist in imagery among adults. The relevance of this research also rests on the assumption that identifiable properties of instructional materials can and do affect the performance of students. The research task is to identify bases for matching the properties of these materials to the mental activities of students who have been identified as imagers or verbalizers. Thus, for example, the evidence might suggest that visualizers (i.e., imagers) would profit most (from the standpoint of acquisition and retention) by the use of material presented pictorially rather than in abstract form. Since imagery as a process is a matter of transformation (encoding) an understanding of its limitations, that is, where it interferes with learning, would be essential for improving instructional strategy. Furthermore, we would expect contextual

constraints to affect pictorial material differently than they would affect verbal material and these effects should interact with individual differences.

Procedure

Several tests were administered to more than three hundred students. There were three spatial relations and perceptual tests for measuring imagery and a test of intelligence to measure the ability to utilize symbolic material in thinking. In addition, the Scholastic Aptitude Test and the Remote Associates Test scores will be available along with other measures of ability to transform stimuli symbolically.

In one experimental task 100 pictorial and verbal stimuli were presented in random order to all Ss tested. Measures of recognition and of recall were obtained. These responses will be scored for differential recall and subjective organization according to whether the bases of retrieval were pictures or words. In a second task, measures were obtained of the rapidity with which half of the Ss learned lists of high, medium, and low imagery (using Paivio's norms) words, and in which the other half of the Ss learned lists of high, medium and low vividness (using Tulving's norms) words. In a third task, the Ss learned to associate a number to either a word or a picture. These conditions were orthogonally crossed with transfer tasks which involved generalized verbal or pictorial stimuli. In a fourth set of tasks the Stroop color-name test and the automatization test, both of which involve competition between verbal and pictorial stimuli, were administered. All data except those obtained in the second task were gathered on all subjects. The second task was divided equally among the Ss; half of

the Ss received lists based on imagery ratings and the other half received lists based on vividness ratings.

The analysis will consist of correlations and, perhaps, a factor analysis of all test scores, to determine whether clear categories of imagers - nonimagers emerge. All data from the experiments will be analyzed via analysis of variance for differences among treatments and interaction of these treatment with individual differences. In addition, individual differences in imagery will be incorporated into the analysis of variance by employing appropriate designs.

Progress

At the time this report is being prepared all phases of test administration and conduct of experiments have been completed. All tests have been scored. There remains the task of putting this data on computer cards, following which the data must be analyzed. Our present plans are to prepare separate reports of several aspects of the study.

References

Broadbent, D. E. Perception and communication. New York: Pergamon, 1958.

Rohwer, W. D. Jr. Images and pictures in children's learning: research results and educational implications. Psychological Bulletin, 1970, 73, 393-403.

Paivio, A. On the functional significance of imagery. Psychological Bulletin, 1970, 73, 385-392.

Presentation Content, Classroom Notes, and Fact Versus Generalization Learning

Study Directors: D. L. Peters and Carl Harris

Purpose

When material is presented to students in classroom situations, both the content and the organization are selected by the instructor. A content analysis of the presentation can provide a description of the objective stimuli. However, as Rothkopf (1968) has indicated, the nominal stimuli and the effective stimuli, i.e., the information as processed by the learner, are not the same. The information actually selected, organized, coded, and reviewed by the student during the study-learning period is likely to be more important to subsequent performance than is the manner of presentation, the content during presentation and the organization of the presentation.

The present study seeks to analyze the relationship between the content and organization of presentation, the content and organization of students' notes (as an indication of what has been selected, coded, and reviewed) and performance on factual versus generalization type test items.

Relevance for Instruction

Frequently, educational research has failed to realize the importance of student instrumental activities in learning. It has assumed that the relationship between what is taught and what is

learned is direct. Variations among students taught with a single procedure is assumed to be error variance. However, under the relatively uncontrolled classroom conditions much of this variance may be accounted for by the type and proficiency of the instrumental activities engaged in by students. Therefore, analysis of the relations among presentation content, learning activities, and criteria will help to define this important area of individual differences.

Procedures

Three types of information (historical, factual but non-historical, and theoretical) were integrated into a single presention of specifiable content and organization. Subjects were pretested on the material and listened to a taped presentation during which they were encouraged to take notes. They were provided with review time and subsequently posttested on the material. Both the pre- and post-tests included factual (requiring information found in the presentation) and generalization (or inference) types of items. The latter required integration or going beyond the information provided. The notes of the subjects were collected to be analyzed for type of content, organization, quantity, and other characteristics. The procedures are designed to allow specification of the presentation, the content and organization of student's notes, and the type of test questions answered.

Progress

The data are collected and are being coded. Technical difficulties with the computer program being utilized for the content analysis of the notes have delayed further analysis to date. This difficulty should be overcome in the near future.

References

Rothkopf, E. Z. Two scientific approaches to the management of instruction. In R. Gagne and W. Gephart (Eds.), Learning research and school subjects. Itasca, Illinois: F. E. Peacock Publishers, 1968.

Note-taking, Rate of Presentation and Immediate Recall

Study Directors: D. L. Peters, Carl Harris and Victor Messier

Purpose

Although generally advocated in education, there is a paucity of systematic research on the effects of note-taking upon the immediate recall of material. Previous research seems to indicate that taking notes while listening either has no effect on immediate recall (Pauk, 1963; McClendon, 1956) or that it is beneficial (Crawford, 1925; McHenry, 1969; Peters and Harris, 1970). However, the effects of note-taking under different presentation rates has not been studied. It could be hypothesized that increasing the rate of presentation decreases the value of taking notes. Additionally, previous research has not been sufficiently controlled to determine if the effects of note taking on subsequent test performance actually involved listening rather than some other aspect of information processing.

The present study attempts to do three things. First it seeks to evaluate the effects of note-taking upon the immediate recall performance of subjects when the presentation rate of the material is varied. Secondly, it seeks to compare the effects of note-taking in "listening" situations with the effects of note-taking in "reading" situations. Lastly, it attempts to determine the relevance of individual differences in oral reading rate and listening efficiency to recall performance of subjects under the different treatment conditions.

Relevance for Instruction

The study of the characteristics of student instrumental activities and the variables affecting the value of these activities for learning will provide a broader understanding of the learning process in the classroom. Taking notes has long been considered a valued behavior on the part of students. The present research should help to clarify the role of this student behavior in learning.

Procedures

Subjects were individually tested for oral reading rate and listening efficiency and assigned to one of two note-taking conditions (no notes or notes encouraged) and one of three presentation conditions (taped normal rate, taped rapid rate, or rapid reading). They were then presented with a 1,613 word passage of scientific material. Following the presentation they were tested on the material they had heard or read.

Progress

The data have been collected and are being analyzed.

References

Crawford, C. C. Some experimental studies of the results of college note-taking. Journal of Educational Research, 1925, 12, 379-86.

McClendon, P. An experimental study of the relationship between the note-taking practices and listening comprehension of college freshmen during expository lectures. Unpublished doctoral dissertation, State University of Iowa, 1956.

McHenry, D. The effect of note-taking activity on listening comprehension in an immediate recall situation. U.S. Air Force Academy Research Report 69-2, 1969.

Pauk, W. Does note-taking interfere with listening comprehension? Journal of Developmental Reading, 1963, 6, 276-278.

Peters, D. & Harris, C. Note-taking and review in reception learning. (Presented in this Annual Report), 1970.

Rote and Conceptual Aspects of Classification Learning

Study Directors: Nicholas M. Sanders and Ovid Tzeng

Purpose

In a previous report to the Advanced Research Projects Agency, Sanders (1970) presented a study of the processes involved in rote-based and concept-based classification tasks. He tentatively concluded 1) that processes in the two types of tasks do differ, 2) that the process leading to the criterion of learning in an initial task established a set or expectancy that a second task would require the same process, and 3) that learners differed in their preferences and/or skills in utilizing the two processes.

The study in progress represents a second step of investigation into the differences involved in rote and concept processes in the learner. Two new problems are studied in the present research. First, if prior experience in a rote or concept task establishes a set to use the same processes in a following task, do the learners all learn the same thing when subsequently presented a common task? Specifically, does prior experience in a rote task lead in a following concept task to accurate memory of the instances presented but no generalization to new instances, while prior experience in a concept task leads to the opposite performances? An answer to this question is essential to further specification of the nature of differences between rote and concept processes.

The second problem is the development of valid measures of learner differences in preference for and/or skill in the two types of processes. In the previous study Sanders (1970) found that learner variations in dogmatism were unrelated to learners' performances in the rote and concept task. Rather than attempting to explore other molar personality or cognitive variables possibly related to rote or concept process preferences, the author chose to develop measures much closer in content and procedure to the experimental tasks. Development of satisfactory measures of individual differences in rote and concept processes is important since interpretation of the previous findings rests on the existence of such individual differences.

Relevance for Instruction

Educational objectives as manifest on many classroom tests most likely call for learning by rote processes. Dates, names, and places, as well as the learning of terms and definitions are examples of knowledges attained rotely. Manifestations of conceptual learning are applications of rules, generalizations, and laws to specific settings not encountered previously. If both types of processes are required in attaining various educational objectives, it becomes an important question as to how the two processes are related.

The previous and present studies should be viewed as initial investigations of the effects of previous experience in learning by rote or conceptual processes on the efficiency and nature of learning in a following task in which rote or concept processes are appropriate. Also, satisfactory identification of individual differences in these processes would allow for appropriate individualization of instruction when tasks clearly require either rote or concept processes.

Though the tasks and procedures used in the present project are considerably removed from the usual instructional setting, they are analogous to instructional procedures utilized in the discovery method, in which the instructor structures a set of concrete experiences with the intent that the learners will induce or discover the underlying principle demonstrated in the various concrete experiences. Therefore, the findings of these studies will be related to that particular instructional method.

Procedure

Data collection is carried out in a laboratory setting with one learner at a time. The laboratory session lasts about one hour, and entails the administration of the individual difference measures, the experimental manipulation of task expectancy through having the learner learn either a rote-based classification or a concept-based classification, and a criterion concept-based classification task. The individual difference measures include an initial task designed to assess the learner's preference for rote or concept process, one task to measure concept process skills, and one measure of rote skills. The experimental tasks are the same as those used by Sanders (1970). The criterion task has two parts; a set of learning instances and a set of test instances. There are sixteen learning instances, after which 32 test instances are presented; half the 32 test instances are the same as the learning instances and half are new. The learner is asked to respond to the test instances by remembering whether he had seen the instance before, and, if so, to recall the label assigned. No feedback is given during the test.

Several analyses are planned. To assess the effects of prior experience in a rote or concept task on the performance in the criterion task, three criterion scores will be used: 1) number correct labels given during the learning stage, 2) number correct identifications of new test instances as not being present in the learning set, and 3) number of correct labels applied to new test instances (generalization of the concept). In addition, the individual difference measures will be analyzed for their intercorrelations and for their correlations with performance in the experimental concept and rote tasks.

Progress

Thirty-four subjects have participated in pilot work designed to refine the individual difference measures and criterion task. The individual difference measures will require further development, though the criterion task now appears to have satisfactory instructions and length.

References

Sanders, N. M. The effects of dogmatism on learning and transfer in concept-based and rote-based classification tasks. Semi-annual report submitted to the Advanced Research Projects Agency, 1970, pp. 63-76.

The Effects of Recall Mode and Recall Interval Expectancies on Recall Performance

Study Director: Paul Weener

Assistant: Samuel Rock

Purpose

The way in which a student actively operates on visual or auditory stimulus materials in a learning task is dependent on his expectancies of when and in what form the information will have to be retrieved. The learner can select from a variety of processing and storage operations depending on his perception of the desired output. Some tasks require the learner to focus on isolated bits of information and may require storage of the presented information over very brief periods of time. Simple rehearsal processes may be adequate to fulfill the requirements of such a task. On the other extreme, some tasks may require the learner to focus on broad, integrative principles which require active structuring and reorganization of the presented stimulus materials and to recall the material months or even years later. Such a task requires active transformational and coding processes which are as yet not well understood, and are quite different from the simpler process of rehearsal and short-term storage.

Relevance For Instruction

The type of anticipated test may influence the nature of what the student will learn. For example, multiple choice tests require recognition rather than recall processes and usually measure recall of isolated units of information rather than information which requires

integration of information from different parts of the stimulus material. This research investigates the effects of student expectations for three rather common methods of information retrieval. The information from research of this nature will be relevant to the concern for optimizing instructional methods for individual learners. Given a desired instructional outcome, one mode of test instruction might lead to better achievement than another mode.

Procedure

Six different experimental conditions were created by combining three different "anticipated recall modes" with two different "anticipated recall intervals." The three recall modes were multiple choice test, essay test, and verbal summary to peer group. The two recall intervals were "immediate" and one week. Groups of six participated in the experiment. Each of the six Ss was presented with a set of instructions which stated that he was to read and study a short article and that this would be followed immediately or one week later with one of three recall modes. The material to be studied was a rather difficult passage dealing with principles governing the development of species.

All Ss were then tested immediately on a multiple choice and an essay test, and returned one week later to take the same tests.

Progress

One hundred ten Ss were run during the Spring term, 1970. The data is presently being analyzed.